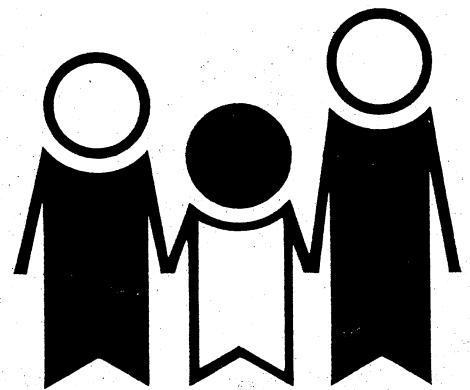


# ADVANCES IN CONSUMER RESEARCH

VOLUME III

---

Edited by  
BEVERLEE B. ANDERSON





# Advances in Consumer Research

## Volume III

Proceedings of Association  
for Consumer Research  
Sixth Annual Conference

Edited by

**Beverlee B. Anderson**

1976 Copyright © Association for Consumer Research

# FORWARD

Since there is otherwise no permanent and public way to acknowledge our appreciation, we are pleased to use this brief space to thank those who contributed to the success of the 1975 ACR conference in Cincinnati.

Obviously, the program committee helped us keep our wits in the midst of a whirlwind of proposals for workshops and special-topic sessions. That its advice was sound is evident in the superlative sessions ultimately chosen for the conference. Not so obvious, but just as important, was the incredible amount and quality of competitive-paper refereeing contributed by literally scores of our colleagues throughout the world. Our pledge of confidentiality prohibits us from identifying them, but hopefully our heartfelt thanks will evade their anonymity.

The officers of ACR helped us in many ways. They militated against those many social murders that typically accompany a conference and its planning. And our President, Jacob Jacoby, had the professional courtesy (and nerve) to let us proceed at will. We don't know how to pay him a higher compliment.

That leaves two debts unrecognized. The conference could not have happened without its participants—whether they were at the speakers' table or in the audience. And this volume would not reflect the professionalism it does were it not for the unyielding standards of Beverlee Anderson.

Finally, while recognizing that dedications can be idolatrous and that melancholy often degenerates into obscenity, let us simply echo the sentiments of hundreds of ACR conferees: "Thanks, Deb."

Jerome B. Kernan  
Stanley D. Shores, Jr.



# PREFACE

Many advances have occurred in the understanding of consumer processes since the founding of the Association for Consumer Research. Each year at the Association meetings, current research in theoretical issues, methodological tools, and applications in the field of consumer behavior are presented and discussed. *Advances in Consumer Research, Vol. III* chronicles the ideas and comments from the Sixth Annual Meeting of the Association, held in Cincinnati, Ohio, October 30 through November 2.

Volume III contains ninety-nine papers and overviews, presented by 138 authors. Ideas of many of the leaders in the field of consumer research are included in the pages of this Volume. New, younger members of the Association, who will be tomorrow's leaders, have contributed their innovative ideas. Many of the contributors to this Volume are affiliated with academic institutions, both from the United States and abroad. Other contributors are associated with government and industrial firms, making this Volume a true reflection of the advances that are being made in consumer research throughout society.

Editing this proceedings has been both a rewarding and frustrating experience. Preparing a proceedings is perhaps analogous to starting a research project, where one is optimistic that the research will be perfect—all endogenous and exogenous variables are controlled; all variables that could cause errors are specified and accounted for. Always, through, the realization comes that errors may be present, regardless of how careful the researcher is. I, too, was optimistic and planned to edit a perfect proceedings. But, as with research, it has become apparent that all errors cannot be controlled. In an attempt to publish the proceedings in a reasonable time period, without forcing the Association into bankruptcy, some adjustments were required. I can honestly say we tried our best for perfection, but please forgive our errors. With slightly mixed emotions I wish to thank Jerry Kernan and Stan Shores for giving me the opportunity to work for six months, editing these proceedings. I hope I have edited a proceedings that meets their expectations.

I want to thank the authors for all their contributions, without which *Advances in Consumer Research, Volume III* would have no substance. The many typists and secretaries who prepared the manuscripts merit my gratitude. The manuscript styling caused many of them strained nerves during the typing process, but I think they will be proud of their contribution to the published work.

Finally, I must mention the excellent work done by Debbie Guethlein, who helped me throughout the duration of this project, Sandy Day, who did a superb job of proof reading many of the manuscripts, and last but not least, Susan Nichols, who did all the correcting and retyping of the manuscripts. It has been a pleasure working with everyone who helped to make this proceedings a volume worthy of the fine material supplied by the members of the Association for Consumer Research.

Beverlee B. Anderson  
Editor





ASSOCIATION FOR CONSUMER RESEARCH

SIXTH ANNUAL CONFERENCE

OCTOBER 30 - NOVEMBER 2, 1975

CINCINNATI, OHIO

CONFERENCE COMMITTEE

Jerome B. Kernan, Co-Chairman  
*University of Cincinnati*

\* \* \*

Beverlee B. Anderson, Proceedings Editor  
*University of Cincinnati*

George E. Brosseau, Jr.  
*National Science Foundation*

Joseph T. Plummer  
*Leo Burnett Advertising*

Robert W. Pratt, Jr.  
*General Electric Company*

Stanley D. Shores, Jr., Co-Chairman  
*Procter & Gamble Company*

\* \* \*

Fred D. Reynolds  
*University of Georgia*

Robert L. Rizek  
*U.S. Department of Agriculture*

Mary Jane Schlinger  
*University of Illinois at Chicago Circle*

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

EXECUTIVE COMMITTEE

Jacob Jacoby, President  
*Purdue University*

David M. Gardner, President-Elect  
*University of Illinois*

William D. Wells, Past President  
*Needham, Harper & Steers Advertising, Inc.*

Robert E. Witt, Newsletter Editor  
*University of Texas*

Jerry C. Olson, Treasurer  
*Pennsylvania State University*

Robert B. Woodruff, Membership  
*University of Tennessee*

Frances M. Magrabi, Publications  
*U.S. Department of Agriculture*

Harold H. Kassarjian, Director  
*UCLA*

M. Venkatesan, Director  
*University of Iowa*

Gerald Zaltman, Director  
*University of Pittsburgh*

Robert W. Pratt, Jr., JCR Policy Board Representative  
*General Electric Company*

ADVISORY COUNCIL

Theodore F. Dunn  
*Benton & Bowles Advertising*

Stephen A. Greyser, *Harvard University and*  
*Marketing Science Institute*

Nancy Harries  
*U.S. Department of Agriculture*

Maureen Kallick  
*University of Michigan*

Fred I. Posner  
*General Foods Corporation*

Jay Schmiedeskamp  
*University of Michigan*

Joseph N. Uhl  
*Purdue University*

LIAISON REPRESENTATIVE FROM NATIONAL SCIENCE FOUNDATION

George E. Brosseau, Jr.



## TABLE OF CONTENTS

<b>PART I -- ADDRESSES</b>		
<i>Consumer Research: Telling It Like It Is</i> ..... 1 Jacob Jacoby	<i>Psychographic Research in a Cross Cultural Nonproduct Setting</i> ..... 98 James F. Engel	
<i>Is All This Advertising Really Necessary?</i> ..... 12 Leo Bogart	<i>Consumer Decisions to Reduce or Stop Using Products and Services: Preliminary Results of a Nationwide Study</i> ..... 102 Douglass K. Hawes, Roger D. Blackwell and W. Wayne Talarzyk	
<i>Interpreting the Polls</i> ..... 17 Seymour Martin Lipset	<i>An Investigation of the Inclusion of the Explicit Ideal Point in the Multi-Attribute Attitude Model</i> ..... 110 Kenneth E. Miller	
<b>PART II -- COMPETITIVE PAPERS</b>		
<i>Attitudinal Influence on Retail Patronage Behavior</i> ..... 24 Masao Nakanishi	<i>An Application of Multidimensional Scaling and Related Techniques to the Evaluation of a New Product Concept</i> ..... 114 Larry Percy	
<i>Customer Stimulation Needs and Innovative Shopping Behavior: The Case of Recycled Urban Places</i> ..... 30 S. L. Grossbart, R. A. Mittelstaedt and S. P. DeVere	<i>A Look at Personality Profiles and the Personality-Attitude-Behavior Link in Predicting Consumer Behavior</i> ..... 119 Larry Percy	
<i>Buyer Behavior Under Conditions of Crowding: An Initial Framework</i> ..... 36 Gilbert D. Harrell and Michael D. Hutt	<i>An Experimental Investigation of Situational Effects on Risk Perception</i> ..... 125 Mark Vincent and William G. Zikmund	
<i>Profiling the Male Fashion Innovator-- Another Step</i> ..... 40 John J. Painter and Kent L. Granzin	<i>Demand Bias in the Assessment of Situational Effects on Buyer Behavior</i> ..... 130 Peter H. Reingen	
<i>Fashion Involvement and Buying Behavior: A Methodological Study</i> ..... 46 Douglas J. Tigert, Lawrence J. Ring and Charles W. King	<i>Dissonance Resolution by Grade School Consumers</i> . 134 David B. Kyner, Jacob Jacoby and Robert W. Chestnut	
<i>An Empirical Evaluation of Comparative Advertising Messages: Subjects' Responses on Perceptual Dimensions</i> ..... 53 R. Dale Wilson	<i>Ideal Point Models of Preference</i> ..... 138 Joel Huber	
<i>Research Methodology and Advertising Substantiation</i> ..... 58 Robert A. Mittelstaedt and Nils-Erik Aaby	<i>An Argument in Support of Ordinary Factor Analysis of Dichotomous Variables</i> ..... 143 Larry Percy	
<i>Consumer Reactions to Comparative Advertising</i> ..... 63 Linda L. Golden	<i>A Multivariate Test of CAD Instrument Construct Validity</i> ..... 149 Michael J. Ryan and Richard C. Becherer	
<i>What is Consumerism?</i> ..... 68 Kent L. Granzin and Gary M. Grikscheit	<i>A Stochastic Inventory Decision Model and the Holding of Family Wealth</i> ..... 155 Ivan F. Beutler	
<i>Consumer Activists: What Makes Them Different?</i> .... 73 Jacques C. Bourgeois and James G. Barnes	<i>A Normative Deficit Model of Consumer Behavior</i> .. 161 Earl W. Morris, Mary Winter and Ivan F. Beutler	
<i>Stochastic Weights in Multiattribute Decision Making</i> ..... 81 Dov Pekelman and Subrata Sen	<i>Comments on a Stochastic Inventory Model and a Normative Deficit Model</i> ..... 166 Richard F. Yalch	
<i>The Role of Random Weights and Reliability in the Assessment of Multiattribute Attitude Models</i> ..... 88 Roger Best, Del I. Hawkins, and Gerald Albaum	<i>Effects of Expectation Creation and Disconfirmation on Belief Elements of Cognitive Structure</i> 168 Jerry C. Olson and Philip Dover	
<i>Operationalizing Risk in Multiattribute Decision Models</i> ..... 92 Bernard Pras and John O. Summers	<i>An Investigation into the Differential Effects of Causally Simple and Complex Attributions</i> ..... 176 Richard Mizerski	

PART II -- (continued)

CONSUMER BEHAVIOR AND PUBLIC HEALTH

*The Identification of Consumer Judgmental Combination Rules: Statistical Prediction vs. Structured Protocol*..... 184  
C. Whan Park and Charles M. Schaninger

*Working Wife vs. Non-Working Wife Families: A Basis for Segmenting Grocery Markets?*..... 191  
Susan P. Douglas

*Spousal Involvement and Empathy in Jointly-Resolved and Authoritatively-Resolved Purchase Subdecision*..... 199  
Alvin C. Burns

*How Will Consumer Education Affect Consumer Behavior?*..... 208  
Paul N. Bloom

*Reflections on Research in Consumer Behavior*... 213  
Johan Arndt

*Behavior Control: Are Consumers Beyond Freedom and Dignity?*..... 222  
Rom J. Markin and Chem L. Narayana

*Correlates of Consumption of Food Nutrients by U. S. Urban Households*..... 229  
Mohamed Abdel-Ghany and Gordon E. Bivens

*The Food Problem and Income Adequacy*..... 238  
Flora Williams

*Attitudes Toward Energy Consumption: Segmenting the Gasoline Market*..... 246  
David J. Barnaby and Richard C. Reizenstein

*Experimentation for Pretesting Public Health Programs: The Case of the Anti-Drug Abuse Campaigns*..... 278  
Michael L. Ray and Scott Ward

*Consumer Response to Seat Belt Use Campaigns and Inducements: Implications for Public Health Strategies*..... 287  
Leon S. Robertson

*Communication Issues in Different Public Health Areas*..... 290  
F. Gerald Kline, Peter V. Miller and Andrew J. Morrison

*The Behavior of the Health Care Consumer: A Selective Review*..... 295  
Lawrence H. Wortzel

*The Role of Mass Communications in Promoting Public Health*..... 302  
Mary Jane Schlinger

PROCESSES OF CONSUMER INFORMATION ACQUISITION

*Pre-Purchase Information Acquisition: Description Of A Process Methodology, Research Paradigm, and Pilot Investigation*..... 306  
Jacob Jacoby, Robert W. Chestnut, Karl C. Weigl and William Fisher

*Patterns of Processing in Consumer Information Acquisition*..... 315  
James R. Bettman and Jacob Jacoby

*Heuristic Search Processes in Decision Making*... 321  
John W. Payne

PART III -- SPECIAL TOPICS

CONSUMER SATISFACTION: A NEGLECTED LINK?

*Consumer Satisfaction as the Ultimate Life Force*..... 252  
Clark Leavitt

*CS/D: The Program Planning and Evaluation Perspective*..... 259  
H. Keith Hunt

*Consumer Satisfaction: A Neglected Link?*..... 261  
Donald J. Hempel and Larry J. Rosenberg

*Consumer Perception of Advertising Effectiveness and Life Style Congruency*..... 328  
D. W. Greeno, G. H. Haines, Jr., and M. S. Sommers

*Consumer Information Acquisition: Public Policy Perspectives*..... 334  
William L. Wilkie

*The Impact of Mobility and Social Integration on Information Seeking*..... 341  
Donald J. Hempel and William J. McEwen

MEASURING CONSUMER COMPLAINTS

*Collecting Comprehensive Consumer Complaint Data by Survey Research*..... 263  
Ralph L. Day and E. Larid Landon, Jr.

ATTRIBUTE MODELS OF CONSUMER CHOICE: RECENT THEORETICAL AND EMPIRICAL DEVELOPMENTS

*Hierarchies in Goods-Characteristics Analysis*.. 348  
Kelvin Lancaster

*Beyond Conjoint Measurement: A Method of Pairwise Trade-Off Analysis*..... 353  
Richard M. Johnson

THE FORMATION OF CONSUMER POLICIES IN THE PUBLIC DOMAIN: CHANGES IN RESEARCH AND EMPHASIS

*Public Demands on Business: A Research Frontier* 269  
William Michael Denney

*When Do Advertisements Mislead the Consumer: An Answer From Experimental Psychology*..... 273  
J. Edward Russo

*The Formation of Consumer Policies in the Public Domain: A Conceptual Overview*..... 276  
Saul Barry Wax

*Attribute Ratings as Predictors of Claimed and Actual Behaviors*..... 359  
Penny Baron and Gerald Eskin

INTERPERSONAL INTERACTION AND PERSUASION PROCESSES

*Influence Processes in Interpersonal Persuasion* 364  
Morris Holbrook and John O'Shaughnessy

PART III -- (continued)

INTERPERSONAL INTERACTION AND PERSUASION...(con't.)

*Situational Influence in Interpersonal Persuasion*..... 370  
Richard J. Lutz and Pradeep Kakkar

*Consumer Decision Making in Naturalistic Settings: Salesman-Prospect Interaction*..... 379  
Richard W. Olshavsky

*Buyer-Seller Interaction: A Conceptual Framework* 382  
Jagdish N. Sheth

*Self-Perception as a Means of Personal Influence: The Foot-In-The-Door Technique*..... 387  
Brian Sternthal, Carol A. Scott and Ruby Roy Dholakia

*Dyadic Interaction: An Exchange Process*..... 394  
David T. Wilson

*Consumer Response to Alternative Selling Strategies: A Field Experiment*..... 398  
Arch G. Woodside and Robert E. Pitts

*Interpersonal Interaction and Persuasion Processes: An Overview*..... 405  
Noel Capon and James Hulbert

CONSUMER RESEARCH ON URBAN TRANSPORTATION PROBLEMS

*Consumer Research in Urban Transportation: Some Methodological Issues*..... 407  
Christopher H. Lovelock

*A Behavioral Travel Demand Model Incorporating Choice Constraints*..... 416  
Wilfred W. Recker and Thomas F. Golob

*A Psychological Model of Travel Mode Selection* 425  
Jagdish N. Sheth

THE GROUP INTERVIEW IN CONSUMER RESEARCH

*Exploratory Group Interview in Consumer Research: A Case Example*..... 431  
Thomas D. Dupont

*Preparing for Group Interviews*..... 434  
Melanie S. Payne

*The Dynamics of the Group Interview*..... 437  
Myril D. Axelrod

*Presearch as Giraffe: An Identity Crisis*..... 442  
Jane Templeton

*What Administrators Should Know About the Group Interview*..... 447  
George J. Szybillo

CONSUMER BEHAVIOR OF THE AGED

*Profiling the Senior Citizen Market*..... 449  
Kenneth L. Bernhardt and Thomas Kinnear

*A Transgenerational Comparison--The Elderly Fashion Consumer*..... 453  
Claude R. Martin, Jr.

*A Descriptive Model of Consumer Choice Processes Among Nursing Home Patients*..... 457  
Steven A. Baumgarten, Tanniru R. Rao and L. Winston Ring

*The Elderly Consumer: One Segment or Many?*.... 463  
Jeffrey G. Towle and Claude R. Martin, Jr.

PART IV -- WORKSHOPS

EXTENDING THE FISHBEIN EXTENDED MODEL

*Conceptual and Operational Issues in the Extended Fishbein Model*..... 469  
Richard J. Lutz

*Fishbein's Subjective Norm: Theoretical Considerations and Empirical Evidence*..... 477  
Myron Glassman and Nancy Fitzhenry

*Toward a Vector Model of Intentions*..... 481  
Olli T. Ahotla

*The Nature of Salient Outcomes and Referents in the Extended Model*..... 485  
Michael J. Ryan and Michael J. Etzel

*Extending the Extended Model: Some Comments*... 491  
Martin Fishbein

MEANING, MEASUREMENT AND OPERATIONAL SIGNIFICANCE OF THE CONCEPT OF LIFE STYLE

*Comment on the Meaning of "Life Style"*..... 498  
William D. Wells

*The Meaning of Life-Style: Sociological and Marketing Perspectives*..... 499  
Paul M. Hirsch

*The Advantages and Disadvantages of the Profile Approach to Analyzing Life Style Data*.. 501  
Stephen C. Cosmas

*Segmentation Analysis--A Tool for Measuring Life Styles*..... 504  
Sunil Mehrotra

*An Exploration of the Role of Family Life Style on Selected Behavior Variables*..... 506  
Lucy Chao Lee

CHILDREN AS CONSUMERS: SOME THEORETICAL PERSPECTIVES

*Children as Consumers: The Need for Multitheoretical Perspectives*..... 508  
Thomas S. Robertson and Shel Feldman

*Children's Social Learning from Television Advertising: Research Evidence on Observational Modeling of Product Consumption*..... 513  
Charles K. Atkin

*Moppets in the Market Place: Evaluating Children's Responses to Television Advertising*. 520  
Barbara R. Fowles

*Visual and Verbal Memory in Children's Product Information Utilization*..... 523  
John R. Rossiter

PART IV -- (continued)

CHILDREN AS CONSUMERS:... (con't.)

<i>Television Commercials as Socializing Agents...</i>	528
Ann H. Beuf	
<i>The Development of Consumer Information- Processing Skills: Contributions from Cognitive Development Theory.....</i>	531
Daniel B. Wackman and Scott Ward	
<i>Cognitive Response to Advertising: The Relation of Child to Adult Models.....</i>	536
Bobby J. Calder, Thomas Robertson and John Rossiter	
Author Index.....	539

ACR PRESIDENTIAL ADDRESS

CONSUMER RESEARCH: TELLING IT LIKE IT IS<sup>1</sup>

Jacob Jacoby  
Purdue University

It was just last week that Cincinnati won the World Series. Among other things, this reminded me of the story about the three umpires who were standing around discussing how they determined whether the pitcher's throw was a ball or a strike. The first umpire said: "Some is balls and some is strikes, and I calls 'em the way they is." The second, perhaps a bit wiser for having had an introductory psychology course, said: "Some is balls and some is strikes, and I calls 'em the way I sees them." The third umpire, possibly an unemployed philosophy Professor, countered with: "Some is balls and some is strikes -- but they ain't nothin' till I calls 'em." A little bit afraid to tread in the area of philosophy, recognizing that this is primarily "the way I sees it," but knowing that there are others here who also see what I see, I've taken the liberty of entitling this talk: "Consumer research: Telling it like it is."<sup>2</sup>

Defining Consumer Behavior<sup>3</sup>

Let me begin by defining what I mean by the terms "consumer research" and "consumer behavior." Several writers -- including some who are members of this Association -- have asserted that consumer behavior is nothing more than a sub-domain of marketing (e.g., Arndt, 1968, p. 5). Others contend that it "is an applied discipline...[which] attempt[s] to solve practical marketing problems" (Engel, Kollat, and Blackwell, 1973, p. 8). Conceivably, authors in other disciplines may claim consumer behavior as their own.

<sup>1</sup>This address is dedicated to my students, past and present, the people who taught me much of what I know.

Apologia. To the captive audience that witnessed my frustrated and somewhat less than adept attempt to compress this paper into a 25 minute address, my apologies. As has been pointed out to me by others, those who study information overload are often the least likely to recognize it in practice.

<sup>2</sup>There have been five previous occasions for outgoing ACR officers to deliver a farewell address. In the days before we had an elected president, Jim Engel and Bob Perloff, the first and second Chairmen of the Advisory Council (then, the governing body of ACR), both spoke about the developing organization and its future. Joel Cohen (1973), our first President, focussed on the potential of ACR to serve as a facilitator of research and, in an all-too-brief section, commented on the trivial nature and poor quality of consumer research, circa 1972. Bob Pratt (1974), our second President, presented a thorough accounting of where this organization had been, what progress had been made, and what progress remained to be made. My immediate predecessor, Bill Wells, chose not to give any address at all. This author confesses that the full wisdom of Wells' decision did not sink in until he started to prepare this address two weeks before it had to be delivered.

<sup>3</sup>This section is based upon Jacoby (1975).

I prefer to view consumer behavior as being independent of any disciplinary orientation. To me, it is a fundamental form of human behavior; it simply exists. Moreover, it will continue to exist regardless of whether or not any discipline makes it the subject of formal inquiry. As I define it (cf. Jacoby, 1975, 1976a, 1976b), consumer behavior is the acquisition, consumption, and disposition of goods, services, time, and ideas by decision making units (e.g., individuals, families, firms, etc.).<sup>4</sup> As such, it represents three broad classes of behaviors (namely, acquisition, consumption, and disposition) directed toward aspects of the environment.

Consumer behavior encompasses much more than just buying products and/or services. For example, acquisition can occur in a wide variety of ways, not all of which involve purchasing and the exchange of money. To illustrate, the sensuous young lady dispensing favors to her 50-year old "uncle" in return for a mink coat, a luxury apartment, and a vacation on the Riviera is, among other things, engaging in a very old form of consumer behavior.

Borrowing the neighbor's rake -- a form of temporary acquisition which does not involve the exchange of money -- can also be considered a form of consumer behavior. One could even argue that the chimpanzee who works for tokens and then exchanges them either for a banana or an opportunity to look through a window into another room is also engaging in consumer behavior (cf. Cowles, 1937; Wolfe, 1936). If you think this is a bit far-fetched, you have only to look into last year's ACR Proceedings to find an empirical investigation which used laboratory animals -- and I don't mean students -- to study economic behavior (Battalio and Kagel, 1975).

Quite clearly, many would end their definition of consumer behavior at this point. The most often cited models of consumer behavior (cf. Engel, Kollat, and Blackwell, 1968, 1973; Howard and Sheth, 1969; Nicosia, 1966; Hansen, 1972) focus predominantly on pre-purchase acquisition; they barely mention or discuss actual consumption (i.e., the decisions and behaviors involved in actually using or consuming the product) and completely ignore disposition. However, consumer behavior also includes these other basic categories of human behavior. Like acquisition, consumption and disposition are both complex decision processes having many facets. Among other things, consumption may be immediate, delayed, or extended through time, and the object of consumption may be entirely consumed (e.g., a cookie) or may remain in complete or partial form after consumption has ceased (e.g., a candy bar wrapper, an old shirt, an auto which is beyond repair). In the latter event, the consumer eventually becomes involved in a decision making and behavioral process regarding whether to throw the object away, give it away, sell it, rent it, convert it to another purpose, etc. Often, the acquisition, consumption, or disposition of one item requires the acquisition, consumption, or disposition of another item (e.g.,

<sup>4</sup>"Behavior" is used here in a general sense to include both overt behavior and related cognitive behavior which may occur prior to, during, and subsequent to overt behavioral acts.

buying a car usually requires that we purchase auto insurance; using a car requires that we purchase gasoline; selling a car usually requires that a new vehicle be acquired, or a new mode of transportation be employed). Thus, consumer behavior often assumes complex overlays of multiple decision making and choice behavior regarding acquisition, consumption, and disposition.

Consumer research, then, is simply research addressed to studying any aspect of consumer behavior. It is not necessarily applied, although it could be and often is. However, it is important to note the growing tendency to consider the study of consumer behavior as a worthy endeavor in its own right (cf. Jacoby, 1969a; Sheth, 1972). In other words, there are also "basic" as well as "applied" consumer researchers. Having now made a distinction between these traditional orientations to research, let me say that the distinction is more arbitrary and artificial than real. Where there are differences between the two, they are more in degree rather than in kind. Applied research almost invariably utilizes basic research concepts and is often concerned with being able to use the obtained information at later points in time (i.e., generalizing). Accordingly, I believe that the issues I am about to raise regarding consumer research are just as relevant for people who call themselves applied as for those who have a more basic orientation.

#### Problems in Consumer Research

In the few passages of his outgoing Presidential Address that were devoted to the subject of consumer research, Cohen (1973, p. 4-5) made the following general observations: "...too much of the research is trivial, both theoretically and for problem solution. Simply put, the quality of our research is not as high as it should be, regardless of purpose." In another paper delivered at that same conference, Kollat, Blackwell, and Engel (1972), based upon reviewing five years worth of the published literature in order to update the Engel, Kollat, and Blackwell (1968, 1973) text, described several of these problems. In most instances, these were the same problems which prevailed when these authors prepared their first edition (Kollat, Engel, and Blackwell, 1970). Perhaps the most telling passage in the Kollat, Blackwell, and Engel (1972, p. 577) paper is the following:

"...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 90% of the findings and lack of findings prove to be wrong..."

Having myself recently prepared a chapter on consumer psychology for the Annual Review of Psychology (Jacoby, 1976b) and considered nearly 1000 articles in the process, my impression is that virtually no progress has taken place in the intervening years for most of the problems that Kollat et al. identified.

I would like to discuss these problems and also raise several new ones for your consideration -- not because they make pleasant reading or listening, but because it is all too apparent that much too large a proportion of the contemporary consumer research literature is not worth the paper it is printed on or the time it takes to read it. Unless we begin to take corrective action soon, we will all drown in a mass of meaningless junk! Let me document this assertion by considering five broad categories of problems: our theories (and comprehensive models), our research methods, our research measures, our statistical techniques, and our subject

matter.<sup>5</sup>

#### Consumer Behavior Theories, Models, and Concepts

The past decade has witnessed an increasing amount of attention devoted to the development, presentation, and discussion of relatively comprehensive theories and models of consumer behavior (Andreason, 1965; Nicosia, 1966; Engel, Kollat, and Blackwell, 1968, 1973; Howard and Sheth, 1969; Hansen, 1972; Markin, 1974). However, Kollat et al. (1972, p. 577) noted that: "These models have had little influence on consumer behavior 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." Unfortunately, not much has changed since then.

Look Ma -- No Theory. Despite the availability of theory and the necessity for theory in any scientific endeavor seeking to extend understanding via empirical research, the impetus and rationale underlying much consumer behavior research seems to rest on little more than the availability of easy-to-use measuring instruments, the existence of more or less willing subject populations, the convenience of the computer, and/or the almost toy-like nature of sophisticated quantitative techniques. Little reliance is placed on theory either to suggest which variables and aspects of consumer behavior are of greatest importance and in need of research, or as a foundation around which to organize and integrate findings. It is still true that nothing is so practical as a good theory. However, while most of us talk a good game about the value and need for theory, it is clear that we would rather be caught dead than using theory.

The Post Hoc, Atheoretic, Shotgun Approach to Conducting Consumer Research. A fundamental problem relating to the neglect of theory and theoretically derived concepts is that the researcher increases the likelihood that he will fail to understand his own data and/or be able to meaningfully interpret and integrate his findings with findings obtained by others. In a set of unpublished working papers now six years old (Jacoby, 1969a, 1969b; as well as in a subsequent empirical investigation, Jacoby, 1971),<sup>6</sup> I referred to the problem as "the athe-

<sup>5</sup>Let me shout it at the outset: MEA CULPA! I have committed many of the sins that I am about to describe. No doubt, I will continue to commit at least some of them long after this address is published and forgotten. There is no one of us without guilt. However, we have to begin casting stones about and break our false idols lest our collective guilt suffocate the periodic airing of our sins and, in so doing, also suffocate the impetus to improve.

I would also like to note at this point that naming names and citing specific articles as illustrations of the problems I am iterating would probably serve few, if any, positive ends. The interested reader has only to examine the articles in our leading journals to find numerous examples of what I mean. On the other hand, and because they may serve a guidance function for some, I have named names and cited specific articles in order to illustrate positive examples addressed to the issue under consideration. It should be noted, however, that citing an article as being positive in one respect usually does not mean that it is void of other deficiencies

Finally, some of the positively cited articles will be my own. I beg the reader's forbearance for the human tendency to be most familiar with and cite one's own work.

<sup>6</sup>Copies of these papers are still available on request.



buying a car usually requires that we purchase auto insurance; using a car requires that we purchase gasoline; selling a car usually requires that a new vehicle be acquired, or a new mode of transportation be employed). Thus, consumer behavior often assumes complex overlays of multiple decision making and choice behavior regarding acquisition, consumption, and disposition.

Consumer research, then, is simply research addressed to studying any aspect of consumer behavior. It is not necessarily applied, although it could be and often is. However, it is important to note the growing tendency to consider the study of consumer behavior as a worthy endeavor in its own right (cf. Jacoby, 1969a; Sheth, 1972). In other words, there are also "basic" as well as "applied" consumer researchers. Having now made a distinction between these traditional orientations to research, let me say that the distinction is more arbitrary and artificial than real. Where there are differences between the two, they are more in degree rather than in kind. Applied research almost invariably utilizes basic research concepts and is often concerned with being able to use the obtained information at later points in time (i.e., generalizing). Accordingly, I believe that the issues I am about to raise regarding consumer research are just as relevant for people who call themselves applied as for those who have a more basic orientation.

#### Problems in Consumer Research

In the few passages of his outgoing Presidential Address that were devoted to the subject of consumer research, Cohen (1973, p. 4-5) made the following general observations: "...too much of the research is trivial, both theoretically and for problem solution. Simply put, the quality of our research is not as high as it should be, regardless of purpose." In another paper delivered at that same conference, Kollat, Blackwell, and Engel (1972), based upon reviewing five years worth of the published literature in order to update the Engel, Kollat, and Blackwell (1968, 1973) text, described several of these problems. In most instances, these were the same problems which prevailed when these authors prepared their first edition (Kollat, Engel, and Blackwell, 1970). Perhaps the most telling passage in the Kollat, Blackwell, and Engel (1972, p. 577) paper is the following:

"...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 90% of the findings and lack of findings prove to be wrong..."

Having myself recently prepared a chapter on consumer psychology for the Annual Review of Psychology (Jacoby, 1976b) and considered nearly 1000 articles in the process, my impression is that virtually no progress has taken place in the intervening years for most of the problems that Kollat et al. identified.

I would like to discuss these problems and also raise several new ones for your consideration -- not because they make pleasant reading or listening, but because it is all too apparent that much too large a proportion of the contemporary consumer research literature is not worth the paper it is printed on or the time it takes to read it. Unless we begin to take corrective action soon, we will all drown in a mass of meaningless junk! Let me document this assertion by considering five broad categories of problems: our theories (and comprehensive models), our research methods, our research measures, our statistical techniques, and our subject

matter.<sup>5</sup>

#### Consumer Behavior Theories, Models, and Concepts

The past decade has witnessed an increasing amount of attention devoted to the development, presentation, and discussion of relatively comprehensive theories and models of consumer behavior (Andreason, 1965; Nicosia, 1966; Engel, Kollat, and Blackwell, 1968, 1973; Howard and Sheth, 1969; Hansen, 1972; Markin, 1974). However, Kollat et al. (1972, p. 577) noted that: "These models have had little influence on consumer behavior 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." Unfortunately, not much has changed since then.

Look Ma -- No Theory. Despite the availability of theory and the necessity for theory in any scientific endeavor seeking to extend understanding via empirical research, the impetus and rationale underlying much consumer behavior research seems to rest on little more than the availability of easy-to-use measuring instruments, the existence of more or less willing subject populations, the convenience of the computer, and/or the almost toy-like nature of sophisticated quantitative techniques. Little reliance is placed on theory either to suggest which variables and aspects of consumer behavior are of greatest importance and in need of research, or as a foundation around which to organize and integrate findings. It is still true that nothing is so practical as a good theory. However, while most of us talk a good game about the value and need for theory, it is clear that we would rather be caught dead than using theory.

The Post Hoc, Atheoretic, Shotgun Approach to Conducting Consumer Research. A fundamental problem relating to the neglect of theory and theoretically derived concepts is that the researcher increases the likelihood that he will fail to understand his own data and/or be able to meaningfully interpret and integrate his findings with findings obtained by others. In a set of unpublished working papers now six years old (Jacoby, 1969a, 1969b; as well as in a subsequent empirical investigation, Jacoby, 1971),<sup>6</sup> I referred to the problem as "the athe-

<sup>5</sup>Let me shout it at the outset: MEA CULPA! I have committed many of the sins that I am about to describe. No doubt, I will continue to commit at least some of them long after this address is published and forgotten. There is no one of us without guilt. However, we have to begin casting stones about and break our false idols lest our collective guilt suffocate the periodic airing of our sins and, in so doing, also suffocate the impetus to improve.

I would also like to note at this point that naming names and citing specific articles as illustrations of the problems I am iterating would probably serve few, if any, positive ends. The interested reader has only to examine the articles in our leading journals to find numerous examples of what I mean. On the other hand, and because they may serve a guidance function for some, I have named names and cited specific articles in order to illustrate positive examples addressed to the issue under consideration. It should be noted, however, that citing an article as being positive in one respect usually does not mean that it is void of other deficiencies

Finally, some of the positively cited articles will be my own. I beg the reader's forbearance for the human tendency to be most familiar with and cite one's own work.

<sup>6</sup>Copies of these papers are still available on request.

oretical shotgun approach" and tried to illustrate its nature by considering empirical attempts to relate personality variables to consumer behavior. Reaching back into ancient history, the most frequently quoted and paraphrased passage from these papers is as follows:

Investigators usually take a general, broad coverage personality inventory and a list of brands, products, or product categories, and attempt to correlate subjects' responses on the inventory with statements of product use or preference. Careful examination reveals that, in most cases, the investigators have operated without the benefit of theory and with no a priori thought directed to how, or especially why, personality should or should not be related to that aspect of consumer behavior being studied. Statistical techniques, usually simple correlation or variants thereof, are applied and anything that turns up looking half-way interesting furnishes the basis for the Discussion section. Skill at post-diction and post hoc interpretation has been demonstrated, but little real understanding has resulted.

These papers went on to advocate and illustrate why it was necessary for consumer researchers to use theoretically derived hypotheses for specifying variables and relationships in advance. That is, they called on consumer researchers (1) to make predictions of differences and no differences, (2) to explain the reasons underlying these predictions, and (3) to do both prior to conducting their research. Look at it this way. You're sitting with a friend watching Pete Rose at bat in the World Series. Pete Rose hits a home run and your friend says: "I knew he was going to hit that home run." In fact, after this bit of post-diction, your friend even continues with a plausible explanation: "He always hits a home run off right-hand pitchers when he holds his feet at approximately a 70° angle to each other and his left foot pointing directly at the pitcher." Think of how much more confident you would have been that what your friend was saying was correct if he made this as a prediction just as Pete Rose was stepping into the batter's box. (Anticipating one of the issues I will raise below, namely replication, think of how much greater confidence you would have if your friend predicted Rose would hit home runs on two subsequent occasions just before Rose actually hit home runs, and also predicted Rose would not hit a home run on eight other instances where Rose did not hit a home run.)

Although considered in the context of relating personality variables to consumer behavior, these working papers also made it clear that almost every aspect of consumer research reflected the atheoretic shotgun approach, particularly when it came to utilizing concepts borrowed from the behavioral sciences. In a word, the problem was pandemic. Yet despite the fact that this passage was later liberally quoted and re-emphasized by such influential writers as Engel, Kollat, and Blackwell (1973, p. 652-53; Kollat, Engel, and Blackwell, 1972, p. 576-77) and Kassarjian (in his frequently cited review of personality research, 1971, p. 416), the impact of these calls for greater reliance on theory and less shotgunning in consumer research has been negligible. Most consumer researchers are still pulling shotgun triggers in the dark.

Concepts Misplaced, or Whoops! Did you Happen to See Where my Concept Went? Even in those instances where consumer researchers seem to be sincerely interested in conducting research based upon a firm conceptual foundation, they sometimes manage to misplace their concepts when it gets down to the nitty gritty. For example,

the author of one recent article states: "...it is imperative that our definition of deception in advertising recognize the interaction of the advertisement with the accumulated beliefs and experience of the consumer." Two paragraphs later he provides a definition which ignores his imperative. He then goes on to propose plans for detecting deception which completely disregard the fact that deception may occur as a function of the prior beliefs of the consumer and not as a function of the ad (or ad campaign) in question.

Another equally frustrating example is provided by those who define brand loyalty as an hypothetical construct predicated upon the cognitive dynamics or the consumer -- and then proceed to base their measure of brand loyalty solely on the buyer's overt behavior. The consumer behavior literature contains an abundance of similar examples of our inability to have our measures of concepts correspond to these concepts.

The "Theory of the Month" Club. Interestingly, however, the failure to use and test existing theories and comprehensive models of consumer behavior has not discouraged some of us from proposing new theories and comprehensive models, thereby providing us with a different kind of problem. Several of our most respected colleagues seem to belong to a sort of "theory of the month" club which somehow requires that they burst forth with new theories periodically and rarely, if ever, bother to provide any original empirical data collected specifically in an attempt to support their theory. Perhaps those with a new theory or model should treat it like a new product: either stand behind it and give it the support it needs (i.e., test it and refine it as necessary) -- or take the damn thing off the market!

Single Shot vs. Programmatic Research. Another theory-related problem evidenced in the contemporary consumer behavior literature is the widespread failure to engage in programmatic research. Judging from the literature published since the inception of ACR, there are fewer than a dozen individuals who have conducted five or more separate investigations in a systematic and sequential fashion which were addressed to providing incremental knowledge regarding the same broad issue. Instead, what we have is a tradition of single shot studies conducted by what one scholar has termed "Zeitgeisters-Shysters" (Denenberg, 1969).

Rarely, however, have single shot investigations answered all questions that need to be answered or made definitive contributions on any subject of importance. Yet many consumer researchers seem to be operating under the illusory and mistaken belief that they are capable of yielding payout of substance and duration. I am not advocating that we do only programmatic research. Having engaged in enough single shot studies myself (e.g., Kyner, Jacoby, and Chestnut, 1975), I full well appreciate the allure, excitement and challenge often inherent in single shot studies and the potential that such studies sometimes have for providing resolution to an applied problem of immediate concern. I also recognize that it is difficult to caution someone in the depth of an infatuation not to be beguiled. However, if we are to deserve the label "serious researcher" and make contributions of substance, it is necessary that a greater proportion of our efforts be programmatic.

Although I consider theory and concepts to be the proper and best starting point for most consumer research, I recognize that some of us consider theory to be irrelevant. So let me now direct some much needed attention to our methods, our measures, our statistics, and our subject matter -- topics which all consumer researchers, whether so-called applied or basic, must share in common.

Verbal Report vs. Actual Behavior. By far, the most prevalent approach to gathering data in consumer research involves eliciting verbal reports from subjects either via an interview or through the use of a self-administered questionnaire. Typically, these verbal reports assess either (1) recall of past events and behavior, (2) current psychological states (including attitudes, preferences, beliefs, statements of intentions to behave and likely reactions to hypothetical events), and/or (3) socio-demographic data. Of the 44 empirical studies in the published Proceedings of last year's conference (Schlinger, 1975), 39 (or 87%) are based principally or entirely on verbal report data collected from respondents. Similarly, of the 36 empirical studies found in the first six issues of the Journal of Consumer Research, 31 (more than 85%) were based primarily or solely on verbal report data. Even if the verbal reports were the best of possible methods, the following observation by Platt (1964, p. 251) would still remain true: "Beware the man of one method or one instrument...he tends to become method-oriented rather than problem-oriented." However, the verbal report is probably not the best of all possible methods. Given the numerous sources of bias in verbal reports and the known and all-too-often demonstrated discrepancy between what people say they do and what they actually do, it is nothing short of amazing that we persist in our slavish reliance on verbal reports as the mainstay of our research.

For the greater part, the problems inherent in the ubiquitous verbal report approach can be organized into one of three broad categories: interviewer error, respondent error, and instrument error. We will here disregard consideration of interviewer errors, since more than 75% of the verbal report studies (or two-thirds of our published empirical effort) are based upon the self-administered questionnaires.

Respondent Error in Verbal Reports. It is exceedingly important to note that verbal report data are predicated upon many untested and, in some cases, invalid assumptions. Many of these are in regard to the respondent. As examples, consider the following assumptions inherent in attempts to elicit recall of factual information: (1) Prior learning (and rehearsal) of the information has actually taken place; that is, something exists in memory to be recalled. (2) Once information is stored in memory, it remains there in accurate and unmodified form. (3) Said information remains equally accessible through time. (4) There are no respondent differences in ability to recall which should be controlled or accounted for. (5) Soliciting a verbal report is a non-reactive act; that is, asking questions of respondents is unlikely to have any impact on them and on their responses.

Analogous assumptions exist with respect to assessing psychological states via verbal reports (e.g., regarding attitudes, preferences, intentions, etc.). For example, in a paper published eight years ago -- which I believe should be required reading for all consumer researchers -- Leo Bogart noted that the simple act of asking the respondent a question often "forces the crystallization and expression of opinions where [previously] there were no more than chaotic swirls of thought" (1967, p. 335). It should be noted that the assumptions underlying recall of factual material are few and simple relative to the assumptions underlying using verbal reports as indicants of psychological states. Perhaps the most effective way to summarize the state of affairs is to simply say that many of the fundamental assumptions which underlie the use of verbal reports are invalid. The reader is asked to persevere regard-

ing the ramifications of this fact.

Instrument Error in Verbal Reports. If these problems are sobering, consider the fact that our paper and pencil instruments (either self-administered questionnaires or interview schedules) often contribute as much or more error than do our interviewers or our respondents. In general, most of our questionnaires and interview schedules are terrible and tend to impair rather than assist us in our efforts to collect valid data. More often than not, we provide respondents with questionnaires which, from their perspective, are ambiguous, intimidating, confusing, and incomprehensible. But questions and questionnaires are easy to prepare, right? Wrong! Preparing a self-administered questionnaire is one of the most difficult steps in the entire research process. Unfortunately, it is commonly the most neglected step. Formulating questions and developing the questionnaire seems like such a simple thing to do that we are usually lulled into a false sense of security. Everyone is assumed to be an expert here. Yet many of us never become aware of the literally hundreds of details that should be attended to in constructing questionnaires (cf. Erdos, 1970; Payne, 1951; Kornhauser and Sheatsley, 1959; Selltitz, Jahoda, Deutch, and Cook, 1959). We simply assume that because we know what we mean by our questions and we comprehend the lay-out and organization of our instrument, data collected using such an instrument are naturally valid. If the data are not valid, then the error is obviously a function of the respondent, not a function of our instrument. The result is that we are often left with what in computer parlance is referred to as GIGO, that is, garbage in-garbage out. In most instances, we ourselves are hardly even cognizant of the fact that this has occurred.

Please don't misinterpret what I am saying. I am NOT suggesting that we do away with verbal reports and self-administered questionnaires. This approach to gathering data is a valid and vital part of our methodological armamentarium. However, if we are to continue placing such great reliance on it, the least we ought to do is clean it up. Too many of us are caught up in the excitement and challenge of research and ignore the basics. One of the things I am most emphatically calling for is for us to get down to these basics, to learn how to formulate questions and structure questionnaires. I care not that a finding is significant, or that the ultimate in statistical analytical techniques have been applied, if the data collection instrument generated invalid data at the outset. Relative to other aspects of conducting research, more time must be devoted to developing and polishing our verbal report instruments. Perhaps if journal editors found it important to require publication of the instrument (or at least the critical questions used), it would stimulate improvement in this area.

Verbal Reports vs. Actual Behavior: Continued. But do we actually have to place slavish reliance on the verbal report? Certainly not! One alternative is to devote less time to studying what people say they do and spend more time examining what it is that they actually do do. In other words, we must begin to place greater emphasis on studying behavior, relative to the amount of effort we place on studying verbal reports regarding behavior. There have been several recent developments in this regard. Since a few of these were discussed at some length yesterday (cf. Jacoby, Chestnut, Weigl, and Fisher, 1975; Payne, 1975) and these remarks will also be available in our Proceedings, I'll not devote additional time to the subject here. Let me simply note that the verbal report and behavioral approaches each have their unique advantages and disadvantages. The optimal procedure would probably involve some combination of both (cf. Wright, 1974). Such an approach is most likely to provide us with a better fix on, and deeper

understanding of our findings.

Consumer Behavior: A Dynamic Process Studied with Static Methods. In addition to the necessity of cleaning up our verbal reports and developing greater attention to alternative approaches, we also need to begin studying consumer behavior (which includes consumer decision making) in terms of the dynamic process that it is. Virtually all consumer researchers tend to consider consumer behavior as a dynamic, decision making, behavioral process. Yet probably 99-plus% of all consumer research conducted to date examines consumer decision making and behavior via static, post hoc methods. Instead of being captured and studied, the dynamic nature of consumer decision making and behavior is squelched and the richness of the process ignored. This is another issue which was treated in detail yesterday, and those interested will be able to pursue this subject in the Proceedings (cf. Jacoby, Chestnut, Weigl, and Fisher, 1975).

Roosters Cause the Sun to Rise. Another methodological issue I would briefly like to mention is the necessity for greater reliance on the experimental method, particularly in those instances where cause-effect assertions are made or alluded to. Examination of our literature reveals a surprising number of instances in which causation is implied or directly claimed on the basis of simple correlation. It bears repeating that no matter how highly correlated the rooster's crow is to the sun rising, the rooster does not cause the sun to rise.

More and Richer Dependent and Independent Variables. A final set of methodological issues I would like to raise at this point -- in part, because they are related to the issue of measurement (particularly validity) to which I will turn next -- concerns the need for research (1) which incorporates measures of a variety of dependent variables, (2) which explores the combined and perhaps interacting impact of a variety of independent variables, and (3) which moves away from using single measures of the same dependent variable. With respect to the former, it is often possible to measure a variety of different dependent variables at little additional cost (e.g., accuracy, decision time, and subjective states in Jacoby, Speller, and Berning, 1974). Unfortunately, opportunities for substantially enhancing understanding through the inclusion of a variety of dependent variables are generally ignored. Equally important, we live in a complex, multivariate world. Studying the impact of one or two variables in isolation would seem to be relatively artificial and inconsequential. In other words, we also need more research which examines the impact of a variety of factors impinging in concert.

It is also all too often true that conclusions are accepted on the basis of a single measure of our dependent variable. The costs involved in incorporating a second or third measure of that same variable are usually negligible, particularly when considered in terms of the increased confidence we could have in both our findings and concepts if we routinely used a variety of indices and found that all (or substantially all) provided the same pattern of results (e.g., Jacoby and Kyner, 1973). This second issue (namely, using multiple measures of the same variable) relates more to the validity of our measures than to our methods, and is elaborated upon below.

#### Consumer Research Measures and Indices

Our Bewildering Array of Definitions. Another problem which Kollat, Blackwell, and Engel (1972) referred to is the "bewildering array of definitions" that we have for many of our central constructs. As one example, at

least 40 different and distinct measures of brand loyalty have been employed in the 300 studies comprising the brand loyalty literature (cf. Jacoby and Chestnut, 1975). Virtually no attempt has been made to weed out the poor measures and identify the good ones. Almost everyone has his own preferred measure and seems to blithely and naively assume that findings from one investigation can easily be compared and integrated with findings from investigations which use other definitions. The same horrendous state of affairs exists with respect to many of our other core concepts and constructs. There are at least four different categories of "innovator" definitions (cf. Kohn and Jacoby, 1973; Robertson, 1971) and three different categories of "opinion leadership" definitions (i.e., self designating, sociometric, and key informant). Each one of these categories can and usually is broken out into several specific forms of operationalizations. As examples, Rogers and Catarno (1962), King and Summers (1970) and Jacoby (1972) all provide different operationalizations of self designating opinion leadership.

More incredible than the sheer number of our measures is the ease with which they are proposed and the uncritical manner in which they are accepted as meaningful indicators. In point of fact, most of our measures are only measures because someone says that they are, not because they have been shown to satisfy the standard measurement criteria of validity, reliability, and sensitivity. Stated somewhat differently, most of our measures are no more sophisticated than first asserting that the number of pebbles a person can count in a ten-minute period is a measure of that person's intelligence; next, conduct a study and find that people who can count many pebbles in ten minutes also tend to eat more; and, finally, conclude from this: people with high intelligence tend to eat more.

Wanted, Desperately: Validity. A core problem is this regard in the issue of validity. Just how valid are our measures? Hardly anyone seems to be interested in finding out. Like our theories and comprehensive models, once proposed, our measures seem to take on an almost sacred and inviolate existence all their own. They are rarely, if ever, examined or questioned. Several basic types of validity exist, although often described with somewhat varying terminology (e.g., American Psychological Association, 1966; Angelmar, Zaltman, and Pinson, 1972; Cronbach, 1960; Heeler and Ray, 1972; Nunnally, 1973). The psychometrician Nunnally, in a highly readable and almost layman-like presentation of the subject, writes of three basic types of validity: content validity (which is generally irrelevant in consumer research), predictive validity, and construct validity. Face validity is a fourth, non-psychometric variety and refers to whether a measure looks like it is measuring what it is supposed to be measuring. Examination of the core<sup>7</sup> consumer behavior journals (Journal of Consumer Research, Journal of Marketing Research, Journal of Marketing, Journal of Applied Psychology, Public Opinion Quarterly, Journal of Consumer Affairs, and Journal of Advertising Research) and conference proceedings (of the Association for Consumer Research, American Marketing Association, and the American Psychological Association's Division of Consumer Psychology) since 1970 -- a body of literature consisting of approximately 1000 published articles -- reveals the following with respect to validity.

Face Validity. First, there are numerous examples of face validity. The measures being used almost always look like they are measuring that which they are supposed to be measuring. However, the overwhelming majority of studies go no further, i.e., provide no empirical support. In other words, face validity is often used as

<sup>7</sup>As considered from my biased perspective, i.e., "as I sees it."

a substitute for construct validity.

Predictive Validity. There are also a sizable number of studies which suggest the existence of predictive validity, that is, the measure in question seems to correlate with measures of other variables as predicted. Unfortunately, many investigators do not seem to recognize that predictive validity provides little, if any, understanding of the reasons for the relationship. One can have a predictive validity coefficient of .99 and still not know why or what it means -- other than the fact that the scores on one measure are highly predictive of scores on a second measure. Indeed, the relationship may even be meaningless. As one concrete example, Heeler and Ray (1972, p. 364) note that Kuehn (1963):

...improved the ability of the Edwards Personal Preference Schedule (EPPS) to predict car ownership. He did it with EPPS scores computed by subtracting "affiliation" scores from "dominance" scores. Such a difference really has no psychological or marketing significance; it is just a mathematical manipulation that happened to work in one situation.

Obviously, high predictive validity doesn't necessarily have to be meaningful.

However, there is one type of predictive validity which receives all too little attention, and that is cross-validity. "Whereas predictive validity is concerned with a single sample, cross validity requires that the effectiveness of the predictor composite be tested on a separate independent sample from the same population" (Raju, Bhagat, and Sheth, 1975, p. 407). It should be obvious that unless we can cross-validate our findings, we may really have no findings at all. Again, examination of the consumer behavior literature reveals few attempts at cross-validation (Kaplan, Szybillo, and Jacoby, 1974; Raju, Bhagat, and Sheth, 1975; Speller, 1973; Wilson, Mathews, and Harvey, 1975).

Construct Validity: A Necessity for Science. From the perspective of science, the most necessary type of validity to establish is construct validity. Examination of the recent published literature indicates that less than 2% of our productivity has been directed toward determining construct validity. A large part of the problem lies in the fact that many researchers appear to naively believe that scientific research is a game played by creating measures and then applying them directly to reality. Although guided by some implicit conceptualization of what it is he is trying to measure, the consumer researcher rarely makes his implicit concepts sufficiently explicit or uses them as a basis for developing operational measures. Yet virtually all contemporary scholars of science generally agree that the concept must precede the measure (e.g., Massaro, 1975, p. 23; Plutchik, 1968, p. 45; Selltitz et al., 1959, p. 146-47).

It is not my intention to get into a lengthy discussion of the nature of scientific research.<sup>8</sup> I simply wish to point out that many of our measures are developed at the whim of a researcher with nary a thought given to whether or not it is meaningfully related to an explicit conceptual statement of the phenomena or variable in question. In most instances, our concepts have no identity apart from the instrument or procedures used to measure them. As a result, it is actually impossible to evaluate our measures. "To be able to judge the relative value of measurements or of operations requires criteria

<sup>8</sup>The interested reader is referred to Chapter 4 in Jacoby and Chestnut (1975) for an extended discussion of these issues.

beyond the operations themselves. If a concept is nothing but an operation, how can we talk about being mistaken or about making errors?" (Plutchik, 1968, p. 47). In other words, clearly articulated concepts (i.e., abstractions regarding reality) must intervene between reality and the measurement of reality.

Probably the most efficient means for establishing construct validity is the Campbell and Fiske (1959) multi-method x multi-trait approach. Despite the fact that numerous articles refer to this approach as something that could or should be applied, considerably less than 1% of our published literature has actually employed this approach for systematically exploring construct validity (Davis, 1971; Jacoby, 1974; Silk, 1971). Yet if we cannot demonstrate that our concepts are valid, how can we continue to act as if the findings based upon measures of these concepts are valid? As Campbell and Fiske (1959, p. 100) note: "Before one can test the relationship between a specific trait and other traits, one must have confidence in one's measure of that trait."

Convergent Validity. One basic and relatively easy to establish component of construct validity is convergent validity. This refers to the degree to which attempts to measure the same concept using two or more different measures yield the same results. Even if there are few full-scale construct validity investigations available, it seems reasonable to expect that we should find many studies to demonstrate convergent validity. After all, and as noted above, many of our core constructs are characterized by numerous and varied operationalizations. Surely, there have been many investigations which have used two or more measures of these constructs, thereby permitting us to examine convergent validity. Examination of the literature reveals that such is not the case. Somewhat incredibly, only two (out of 300) published studies exist which administered three or more brand loyalty measures concurrently to the same group of subjects, thereby permitting an examination of how these measures interrelated. Our other core constructs fare equally poorly. Data that are available often indicate that different measures of the same construct provide different results (e.g., Kohn and Jacoby, 1973). Given that we cannot demonstrate adequate convergent validity, it should be screamingly obvious that we have no basis for comparing findings from different studies and making generalizations using such a data base. What we urgently need is more widespread use of multiple measures so that we can begin the relatively simple job of assessing convergent validity. We are being strangled by our bad measures. Let's identify and get rid of them.

Reliability. Another fundamental problem with consumer behavior measures is that data regarding their reliability, particularly test-retest reliability, are rarely provided. As an illustration, only a single study appears in the entire 300 item brand loyalty literature which measures the test-retest reliability of a set of brand loyalty measures. A similar state of affairs exists with respect to indices of other core constructs. In particular, consider the case of the test-retest reliability of recall data. In the entire literature on the use of recall data in advertising -- and I suspect that this takes into account several thousand studies -- only two published articles can be found which provide data on the test-retest reliability of recall data (Clancy and Kueskin, 1971; Young, 1972). Alarming, one of these authors (Young, 1972, p. 7) notes that results obtained in ten retests were the same as those in the initial test in only 50% of the cases. Assuming we were ill and actually had a body temperature of 103° Fahrenheit, how many of us would feel comfortable using a thermometer if, with no actual change in our body temperature, this thermometer gave us readings of 97.0°

100.6°, 98.6°, and 104.4°, all within the space of one 15-minute period. Yet we persistently employ indices of unknown reliability to study consumer purchase decisions and behavior. More sobering, we often develop expensive nationwide promotional strategies and wide-ranging public policies based upon findings derived from using such indices. Obviously, reliability should not only be a concern in Ph.D. dissertations and M.S. theses.

Open Publication Tradition. Let me digress for a moment -- because this seems to be as good a point as any -- to briefly touch upon my use of and stress upon the words "published literature." No doubt, work has been conducted by and for industry which addresses many of these fundamental issues. Much of this work is of high quality. Rarely, however, are the findings from these investigations permitted to enter the published literature. Although there are several reasons for this, a dominant reason is that industry is phobic. Firms are afraid that by permitting such data to be published, they will be giving up trade secrets and competitive advantages. I submit that, in the long run, industry probably has more to gain than lose by permitting this material to surface. No single firm has the resources necessary to make progress along all, or even a sizable proportion of the important research fronts. Contributing to the basic fund of knowledge would yield dividends to all.

Replication. There is a strong necessity for us to replicate our findings using different subject populations, test products, etc. Being able to predict that Johnny Bench will hit a home run on one occasion is not as impressive as being able to accurately predict on two or more occasions. The name of the game is confidence in our findings.

Measurement Based on House-of-Cards Assumption. Another problem which makes its appearance in the literature with alarming frequency is the tendency to have one's measures (or proposed measures) rest upon an intertwined series of untested and sometimes unverifiable assumptions so that the measures used are sometimes 5 or even 15 steps removed from the phenomenon of interest. The article on deceptive advertising noted earlier provides a good case in point. Interpreting data collected via such measures of measurement systems represents a form of specious logic. In such cases, if a single one of the many assumptions is rendered invalid, the entire measurement system must necessarily come cascading downward. However, perhaps there is a positive side to this problem in that it indicates consumer researchers are at least beginning to recognize that their measurements are predicated upon basic assumptions. Stating these assumptions in clear and explicit detail is a necessary and important step before meaningful progress can be made.

The Folly of Single Indicators. A final measurement problem I would like to note is perhaps most easily illustrated by posing the following question: "How many of us would feel comfortable having our intelligence assessed on the basis of our response to a single question?" Believe it or not, that's exactly the kind of thing we do in consumer research. As examples, brand loyalty is often measured by the response to a single question. The same is true with respect to virtually all of our other core constructs. Just a few months ago I came across an exceedingly expensive, large scale, multinational study of consumer information seeking which assessed opinion leadership on the basis of each subject's response to a single question. Examination of our literature reveals hundreds of instances in which responses to a single question suffices to establish the person's level on the variable of interest and then serves as the basis for extensive analysis and en-

tire articles.

Just as is true of such constructs as personality and intelligence, most core concepts in consumer research (e.g., opinion leadership, brand loyalty, innovation proneness, shopping proneness, etc.) are multifaceted and complex. Intelligence and personality are generally measured through the use of a battery of different test items and methods. Even single personality traits are typically assessed by 30 or 40 item inventories. Given the complexity of our subject matter, what makes us think that we can use responses to single items (or even to two or three items) as measures of these constructs, then relate these scores to a host of other variables, arrive at conclusions based upon such an investigation, and get away calling what we have done "science"?

#### Statistics in Consumer Research

Let me now turn to a consideration of the manner in which we use statistics to analyze our data. In general, this is the area where we have the fewest number of problems and, in recent years, probably the greatest number of advances. However, as I see it, we still do have three major problems which I will call "number crunching", "using calipers to measure melting marshmallows", and "static state statistics".

Number Crunching. I have finally reached the point where I am no longer automatically impressed by the use of high-powered and sophisticated statistics. Why? Because too often the use of these statistics appears not to be accompanied by the use of another high-powered and sophisticated tool, namely the brain. For example, what does it really mean when the fourteenth canonical root is highly significant and shows that a set of predictors including size of house, purchase frequency of cake mix, and number of times you brush your teeth per day is related to age of oldest child living at home, laundry detergent preference, and frequency of extra-marital relations? Given the penchant that some have for coming up with brilliant interpretations of such findings, let me hasten to add that my question was simply rhetorical. Of course, this particular mindless application of high-powered statistics is only a way-out example -- or is it? A critical examination of the recent consumer research literature will reveal many more instances of such mindless and mindblowing applications.

Multilayered Madness. In its most sophisticated (a word which, it should be remembered, derives from sophism) form, number crunching involves the multilayering of statistical techniques so that the output from one analysis provides the input for the next analysis. Sometimes, this statistical version of musical chairs involves five to ten different techniques used in series. Again, given the nature of the data collected in the first place, what does the final output actually mean?

Measuring Giant Icebergs in Millimeters and Using Calipers to Measure Melting Marshmallows. Perhaps what is most surprising about this number crunching is the fact that the data being crunched are usually exceedingly crude and coarse to begin with. As already noted, the large majority of our data are collected using the self-administered questionnaire. Yet many consumer researchers don't have the foggiest idea about what the basic do's and don'ts are when it comes to questionnaire construction. Consider also the fact that the reliability and validity of the data we collect are often assumed, not demonstrated. Finally, also consider the fact that trying to measure diffuse, complex, and dynamic variables such as personality, attitudes, motives, brand loyalty, information seeking, etc. may be like trying to measure melting marshmallows with vernier calipers.

In other words, what are we doing working three and four digits to the right of the decimal point? What kind of phenomena, measures, and data do we really have that we are being so precise in our statistical analyses? I submit that our statistical methods are already too sophisticated for the kinds of data we collect. What we need are substantial developments in both our methodology (particularly in regard to questionnaire construction) and in the psychometric quality of our measures (particularly in regard to validity and reliability) before use of the high-powered statistics can be justified in many of the instances where they are now being routinely applied.

Static States Statistics. There is one area, however, in which our statistics can use some improving. By and large, most of our statistics are appropriate only for use with data which are collected using our traditional cross-sectional, static methodologies. However, just as we have a need for the further development of dynamic methodologies, we need the development of statistics for analyzing data collected using such methods. That is, we need statistics which do not force dynamic process data to be reduced to static state representations. To a certain extent, trend analysis, and cross-lagged correlations can and have been used in this manner. However, our repertoire of statistical techniques for handling dynamic data needs to be expanded, either by borrowing from disciplines accustomed to dealing with dynamic data, or through the creative efforts of statisticians working within the consumer research domain.

#### Consumer Research Subject Matter

A final set of consumer research problems I would like to touch upon concerns our subject matter. Joel Cohen called much of it "trivial". In too many ways, he is still right.

Systematically Exploring the Varieties of Acquisition. To begin with, most definitions of consumer behavior tend to shackle us by confining our attention to the purchase of products and services. Aside from the fact that purchase can itself take a variety of forms (e.g., buying at list price, bargaining, bidding at auction), purchase is but one form of acquisition. There are many others. Receiving something as a gift, or in trade, or on a loan basis are three such examples. Each of these can have important economic, sociological, and psychological dynamics and consequences different from purchase. For example, on an aggregate level, if one million more Americans this year than last suddenly decided to borrow their neighbor's rake to handle their fall leaf problems, the impact on the rake industry could be enormous. For that matter, what are the dynamics underlying being a borrower or being a lender? What are the dynamics underlying giving or receiving a gift (cf. Hart, 1974; Weigl, 1975)? Hardly any published data exist regarding these facets of acquisition. Obviously, one thing we must do is systematically explore the realm of consumer acquisition decisions and behavior.

Putting Consumption Back into Consumer Behavior. Although considerable work has been done on consumption, particularly by the home economists, this fact is not adequately reflected in the predominant theories and textbooks of consumer behavior. The work dealing with consumption itself must be given greater salience and more tightly integrated with the existing consumer behavior literature.

And What about Disposition? The third major facet of consumer behavior, namely disposition, appears to have been completely neglected. This unfortunate state of affairs should be rectified for at least four reasons. First, from a purely scholarly perspective, disposition

decisions deserve to be studied in their own right. The scientific approach requires that we study all aspects of a phenomenon, not just part of it. This is particularly important in this instance, since many disposition decisions have significant economic consequences for both the individual and society. Some disposition decisions (e.g., when and how to properly dispose of unused or outdated prescription drugs) may even have important health and safety ramifications. Second, on more practical grounds, much consumer behavior seems to be cyclical and a variety of marketing implications would most likely be forthcoming from an understanding of the disposition subprocess. Third, we are entering an age of relative scarcity in which we can no longer afford the luxury of squandering our natural resources. Understanding consumer disposition decisions and behavior is a necessary (and perhaps even logically prerequisite) element in any conservationist orientation. Finally, the study of consumer disposition could conceivably provide us with new "unobtrusive" (cf. Webb, Campbell, Schwartz, and Sechrest, 1966) macro indicators -- both leading and trailing -- of economic trends and the state of consumer attitudes and expectations. An empirical and taxonomic start toward exploring consumer disposition has recently been made (Jacoby, Berning, and Dietvorst, 1975).

Consumption and Production. Not only does the definition of consumer behavior have to be expanded and its various facets studied, but the relationship between consumption and production should be explored. As implied by the "leaf rake" example above, consumption and production are integrally related. Studies are needed which examine this interrelationship by considering both domains simultaneously.

Addressing Important Social Issues. Much of our subject matter is obviously a function of the pressing social issues which confront us. Or is it? Probably the most significant and potentially overwhelming problem that we as a nation -- and, indeed, the entire world -- have ever confronted is the emerging energy crisis. This problem dwarfs the Viet Nam war in its heyday, the Arab-Israeli situation, our economic stability, misleading advertising, nutrition labeling, and any other problem you can think of. These other problems are all pimples compared to the rogue elephant that is the emerging energy crisis. Yet the total contribution on this subject appearing in the consumer literature is fewer than five empirical and non-empirical papers. Even in those subject areas where we have supposedly been devoting attention, our record is not much better. Regardless of quality, how much empirical work, as opposed to rhetoric, has actually been addressed to the issues of consumer behavior and the elderly, product safety, deceptive and misleading advertising, nutritional labeling, etc., etc? In general, far fewer than ten published studies exist on each of these topics. As Cohen noted, we need to stop toying with the trivial and start addressing that which is significant.

#### Exhortation

This compendium (summarized for students in Table 1) is by no means an exhaustive iteration of all of the problems in and confronting consumer research. Among others, I have not touched upon the widespread tendency to over-generalize from our results, our relative inattention to cross-cultural comparisons, and the numerous avoidable or controllable problems which crop up in regard to the use of experimental designs in our research (cf. Campbell and Stanley, 1963; Rosenthal and Rosnow, 1969). The compendium does, however, cover what I view to be the most frequently occurring and severe problems which confront us.

TABLE 1

An Incomplete Compendium of Major<sup>1</sup> Problems in  
Consumer ResearchTheories, Models, and Concepts

- Theory unused  
 as a basis for guiding research  
 -lack of a priori specification of variables  
 -lack of a priori specification of relationships  
 -reliance on shotgunning  
 as a basis for integrating research  
 as a basis for generalizing from research
- Theory noted, then ignored  
 in going from concept to measurement  
 in interpreting data obtained
- Theory proliferation  
 untested "theory of the month" club contributions  
 Zeitgeistig (single-shot vs. programmatic research;  
 see Denenberg's 1969 index)

Methods

- Verbal reports: underlying assumptions and problems  
 respondent error  
 -in assessing prior events via recall  
 -in assessing current or anticipated psychological states  
 instrument error: our soft vulnerable underbelly  
 -errors too numerous to cite individually  
 -must return to the basics
- Slavish reliance on the verbal report  
 failure to develop and utilize alternatives
- Consumer behavior as a static vs. dynamic process
- Correlation ≠ causation  
 necessity for greater reliance on experimental method
- Our dependent and independent variables  
 need to simultaneously employ a variety of dependent variables  
 need to simultaneously explore a variety of independent variables  
 need to simultaneously utilize several indicants of the same dependent variable

Measures

- Our operational definitions: a bewildering array
- Validity  
 predictive validity  
 -cross-validation  
 construct validity - essential and missing  
 -full test (e.g., multimethod x multitrait approach)  
 -convergent validity - relatively easy but generally ignored
- Reliability  
 test-retest reliability
- Replication
- Measurement based on house-of-cards assumptions
- The folly of single indicants

Statistics

- Number-crunching  
 multilayered number-crunching
- Measuring giant icebergs in millimeters  
 measuring melting marshmallows with vernier calipers
- Static-state statistics

<sup>1</sup>"as I sees it"Content

- Systematically exploring acquisition  
 Integrating consumption into consumer research  
 What about disposition?  
 Consumption and production  
 Addressing important issues  
 the emerging energy crisis  
 everything else

Most of these have been previously discussed in print by one or more of us within the consumer research community. The problems are serious and bear periodic repeating. Some are easier to attend to than others. Hopefully, sensitization will produce awareness which, in turn, will provide the impetus for change.

Quite clearly, I think it's important to know that we don't know -- important so that we don't delude ourselves and others about the quality of our research and validity of our findings as providing sound bases upon which to make decisions of consequence. It is also important to recognize that we are in the midst of a consumer research information explosion and unless we take corrective action soon, we stand to become immersed in a quagmire from which it is already becoming increasingly difficult to extricate ourselves. Perhaps one of the things we most need to learn is that we must stop letting our existing methods and tools dictate and shackle our research. They are no substitute for using our heads. The brain is still the most important tool we have and its use should precede more than succeed the collection of data.

Because I have chosen to focus on our problems, the tone of this address has been rather negative. However, I would like to conclude on what I believe is a very legitimate positive note. Almost every one of the problems noted provides us with numerous opportunities to make meaningful contributions. Simply establishing the validity of a single one of our core constructs and shucking off our poor measures of this construct will require a substantial effort. Consider, also, the need to develop a process technology (incorporating appropriate process methods and statistics) for examining consumer behavior in terms of the dynamic, ongoing phenomenon that it is. As another example, we have need for reviews which not only summarize, but also critically evaluate the empirical evidence bearing on the adequacy of our concepts and measures. Numerous other opportunities become apparent from a consideration of our problems.

It is important to periodically take stock of where we are. However, it is probably more important that we give more than just lip service to these issues; we must begin doing something about them. The time is already overdue.

Having started with a story, let me end with one. Having just met each other for the first time, a young man and young woman are standing together in quiet conversation at a cocktail party. Without any prior indication, the young woman propositions the young man. "Fine. My place or yours?" came his reply. "Well, if it's going to be such a hassle, let's forget about it," said she. The point: it's really not such a hassle to improve consumer research. So why don't we get it on?<sup>9</sup>

<sup>9</sup>This address has hopefully been written so that advanced undergraduates would be able to comprehend the problems being described. If they do, and if they then begin to ask incisive and critical questions of their professors (e.g., "What does this really mean?" and "Why?"), progress is likely to come all the more rapidly. Let's hope so.



## References

- American Psychological Association, Standards for Educational and Psychological Tests and Manuals (Washington: American Psychological Association, 1966).
- Alan R. Andreason, "Attitudes and Consumer Behavior: A Decision model," in L. Preston, ed., New Research in Marketing (Berkeley, Calif.: Institute of Business and Economic Research, University of California, 1965), 1-16.
- Reinhard Angelmar, Gerald Zaltman, and Christian Pinson, "An Examination of Concept Validity," in M. Venkatesan, ed., Proceedings of the Third Annual Conference of the Association for Consumer Research, (1972), 586-93.
- Johan Arndt, ed., Insights into Consumer Behavior (Boston, Mass.: Allyn and Bacon, Inc., 1968).
- Raymond C. Battalio and John H. Kagel, "Experimental Studies of Consumer Demand Behavior: Towards a Technology of Making the Slutsky-Hicks Theory Technologically Applicable to Individual Behavior," in M. J. Schlinger, ed., Advances in Consumer Research: Vol. 2, (Proceedings of the Association for Consumer Research, Chicago: University of Illinois, 1975), 657-70.
- Leo Bogart, "No Opinion, Don't Know, and Maybe No Answer," Public Opinion Quarterly, 31(Fall, 1967), 331-45.
- Donald T. Campbell and Donald W. Fiske, "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," Psychological Bulletin, 56(1959), 81-105.
- Donald T. Campbell and Julian C. Stanley, Experimental and Quasi-Experimental Designs for Research (Chicago: Rand McNally, 1963).
- Kevin J. Clancy and David M. Keskis, "T.V. Commercial Recall Correlates," Journal of Advertising Research, 11(April, 1971), 18-20.
- Joel Cohen, "Presidential Address," untitled, Association for Consumer Research Newsletter, 3(January, 1973), 3-5.
- John T. Cowles, "Food Tokens as Incentives for Learning by Chimpanzees," Comparative Psychological Monographs, 4(5, 1937).
- Lee J. Cronbach, Essentials of Psychological Testing, 2nd ed. (New York: Harper & Bros., 1960).
- Harry L. Davis, "Measurement of Husband-Wife Influence in Consumer Purchase Decisions," Journal of Marketing Research, 8(August, 1971), 305-12.
- Victor H. Denenberg, "Proximites A. Zeitgeister, B.S., M.S., PhONY," Psychology Today, 3(June, 1969), 50.
- James F. Engel, David T. Kollat, and Roger D. Blackwell, Consumer Behavior (New York: Holt, Rinehart & Winston, 1968).
- James F. Engel, David T. Kollat, and Roger D. Blackwell, Consumer Behavior, 2nd ed. (New York: Holt, Rinehart & Winston, 1973).
- Paul L. Erdos, Professional Mail Surveys (New York: McGraw-Hill, 1970).
- Flemming Hansen, Consumer Choice Behavior: A Cognitive Theory (New York: Free Press, 1972).
- Edward W. Hart, Jr., "Consumer Risk-Taking for Self and for Spouse," unpublished Ph.D. dissertation (Purdue University, 1974).
- Roger M. Heeler and Michael L. Ray, "Measure Validation in Marketing," Journal of Marketing Research, 9(November, 1972), 361-70.
- John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior (New York: Wiley, 1969).
- Jacob Jacoby, "Toward Defining Consumer Psychology: One Psychologist's Views," Purdue Papers in Consumer Psychology, No. 101, 1969. Paper presented at the 77th Annual Convention of the American Psychological Association, Washington, D.C., (1969a).
- Jacob Jacoby, "Personality and Consumer Behavior: How NOT to Find Relationships," Purdue Papers in Consumer Psychology, No. 102, (1969b).
- Jacob Jacoby, "Personality and Innovation Proneness," Journal of Marketing Research, 8(May, 1971), 244-47.
- Jacob Jacoby, "Opinion Leadership and Innovativeness: Overlap and Validity," in M. Venkatesan, ed., Proceedings of the Third Annual Conference of the Association for Consumer Research, (1972), 632-49.
- Jacob Jacoby, "The Construct Validity of Opinion Leadership," Public Opinion Quarterly, 38(Spring, 1974), 81-89.
- Jacob Jacoby, "Consumer Psychology as a Social Psychological Sphere of Action," American Psychologist, 30 (October, 1975), 977-87.
- Jacob Jacoby, "Consumer and Industrial Psychology: Prospects for Theory Corroboration and Mutual Contribution," in M. D. Dunnette, ed., The Handbook of Industrial and Organizational Psychology, (Chicago: Rand McNally, 1976a).
- Jacob Jacoby, "Consumer Psychology: An Octennium," in P. Mussen and M. Rosenzweig, eds., Annual Review of Psychology, 27(1976b), 331-58.
- Jacob Jacoby, Carol K. Berning, and Thomas Dietvorst, "What about Disposition?" Purdue Papers in Consumer Psychology, No. 152, (1975).
- Jacob Jacoby and Robert W. Chestnut, "Brand Loyalty Measurement: A Critical Review," monograph submitted for publication, (1975).
- Jacob Jacoby, Robert W. Chestnut, Karl Weigl, and William Fisher, "Pre-Purchase Information Acquisition: Description of a Process Methodology, Research Paradigm, and Pilot Investigation," in B. B. Anderson, ed., Advances in Consumer Research: Vol. 3, (Proceedings of the Sixth Annual Conference of the Association of Consumer Research, Cincinnati, Ohio, October 30-November 2, 1975).
- Jacob Jacoby and David B. Kyner, "Brand Loyalty vs. Repeat Purchasing Behavior," Journal of Marketing Research, 10(February, 1973), 1-9.
- Jacob Jacoby, Donald E. Speller, and Carol Kohn Berning, "Brand Choice Behavior as a Function of Information Load: Replication and Extension," Journal of Consumer Research, 1(June, 1974), 33-42.
- Leon B. Kaplan, George J. Szybillo, and Jacob Jacoby, "Components of Perceived Risk in Product Purchase: A Cross-Validation," Journal of Applied Psychology, 59 (June, 1974), 287-91.

- Harold H. Kassarian, "Personality and Consumer Behavior: A Review," Journal of Marketing Research, 8(November, 1971), 409-18.
- Charles W. King and John O. Summers, "Overlap of Opinion Leadership Across Consumer Product Categories," Journal of Marketing Research, 7(February, 1970), 43-50.
- Carol A. Kohn and Jacob Jacoby, "Operationally Defining the Consumer Innovator," Proceedings, 81st Annual Convention of the American Psychological Association, 8(2, 1973), 837-38.
- David T. Kollat, Roger D. Blackwell, and James F. Engel, "The Current Status of Consumer Behavior Research: Development During the 1968-1972 Period," in M. Venkatesan, ed., Proceedings of the Third Annual Conference of the Association for Consumer Research, (1972), 576-85.
- David T. Kollat, James F. Engel, and Roger D. Blackwell, "Current Problems in Consumer Behavior Research," Journal of Marketing Research, 7(August, 1970), 327-32.
- Arthur Kornhauser and Paul B. Sheatsley, "Questionnaire Construction and Interview Procedure," in C. Selltitz, M. Jahoda, M. Deutsch, and S. W. Cook, eds., Research Methods in Social Relations (New York: Henry Holt & Co., 1959), 546-87.
- Alfred A. Kuehn, "Demonstration of a Relationship between Psychological Factors and Brand Choice," Journal of Business, 36(April, 1963), 237-41.
- David B. Kyner, Jacob Jacoby, and Robert W. Chestnut, "Dissonance Resolution by Grade School Consumers," in B. B. Anderson, ed., Advances in Consumer Research: Vol. 3, (Proceedings of the Sixth Annual Conference of the Association of Consumer Research, Cincinnati, Ohio, October 30-November 2, 1975).
- Ron J. Markin, Consumer Behavior: A Cognitive Orientation (New York: Macmillan Publishing Co., 1974).
- Dominic W. Massaro, Experimental Psychology and Information Processing (Chicago: Rand-McNally, 1975).
- Francesco Nicosia, Consumer Decision Processes (Englewood Cliffs, N. J.: Prentice-Hall, 1966).
- Jum C. Nunnally, Psychometric Theory (New York: McGraw-Hill, 1973).
- John W. Payne, "Heuristic Search Processes in Decision Making," in B. B. Anderson, ed., Advances in Consumer Research: Vol. 3, (Proceedings of the Sixth Annual Conference of the Association for Consumer Research, Cincinnati, Ohio, October 30-November 2, 1975).
- Stanley L. Payne, The Art of Asking Questions (Princeton, N. J.: Princeton University Press, 1951).
- John R. Platt, "Strong Inference," Science, 146(1964), 347-53.
- Robert Plutchik, Foundations of Experimental Research (New York: Harper & Row, 1968).
- Robert W. Pratt, Jr., "ACR: A Perspective," in S. Ward and P. L. Wright, eds., Advances in Consumer Research: Vol. 1, (Urbana, Illinois: Association for Consumer Research, 1974), 1-8.
- P. S. Raju, Rabi S. Bhagat, and Jagdish N. Sheth, "Predictive Validation and Cross-Validation of the Fishbein, Rosenberg, and Sheth Models of Attitudes," in M. J. Schlinger, ed., Advances in Consumer Research, Vol. 2, (Proceedings of the Fifth Annual Conference of the Association for Consumer Research, Chicago: University of Illinois, 1975), 405-25.
- Thomas S. Robertson, Innovative Behavior and Communication (New York: Holt, Rinehart & Winston, 1971).
- Everett M. Rogers and David G. Cartano, "Methods of Measuring Opinion Leadership," Public Opinion Quarterly, 26(Fall, 1962), 435-41.
- Robert Rosenthal and Ralph L. Rosnow, eds., Artifact in Behavioral Research (New York: Academic Press, 1969).
- Mary Jane Schlinger, ed., Advances in Consumer Research: Vol. 2, Proceedings of the Association for Consumer Research, (Chicago: University of Illinois, 1975).
- Claire Selltitz, Marie Jahoda, Morton Deutsch, and Stuart W. Cook, Research Methods in Social Relations (New York: Henry Holt & Co., 1959).
- Jagdish N. Sheth, "The Future of Buyer Behavior Theory," in M. Venkatesan, ed., Proceedings of the Third Annual Conference of the Association for Consumer Research, (1972), 562-75.
- Alvin J. Silk, "Response Set and the Measurement of Self-Designated Opinion Leadership," Public Opinion Quarterly, 35(Fall, 1971), 383-97.
- Donald E. Speller, "Attitudes and Intentions as Predictors of Purchase: A Cross-Validation," Proceedings, 81st Annual Convention of the American Psychological Association, 8(2, 1973), 825-26.
- Eugene J. Webb, Donald T. Campbell, Richard D. Schwartz, and Lee Sechrest, Unobtrusive Measures: Non-Reactive Research in the Social Sciences (Chicago: Rand-McNally, 1966).
- Karl Weigl, "Perceived Risk and Information Search in a Gift Buying Situation," unpublished M.S. thesis (Purdue University, 1975).
- David T. Wilson, H. Lee Mathews, and James W. Harvey, "An Empirical Test of the Fishbein Behavioral Intention Model," Journal of Consumer Research, 1(March, 1975), 39-48.
- John B. Wolfe, "Effectiveness of Token-Rewards for Chimpanzees," Comparative Psychological Monographs, 12(5, 1936).
- Peter L. Wright, "Research Orientations for Analyzing Consumer Judgement Processes," in S. Ward and P. L. Wright, eds., Advances in Consumer Research: Vol. 1, (Urbana, Illinois: Association for Consumer Research, 1974) 268-79.
- Shirley Young, "Copy Testing Without Magic Numbers," Journal of Advertising Research, 12(February, 1972), 3-12.

P.S. Because her name appears nowhere else in print in connection with this conference, on behalf of the Association, I would like to express our warm and sincere appreciation to Miss Deborah Guethlein (Jerry Kernan's secretary) for all she did to make the 1975 Conference the success that it was.

## IS ALL THIS ADVERTISING REALLY NECESSARY?

Leo Bogart, Newspaper Advertising Bureau, Inc.

My off-the-cuff response to the question, "Is all this \$30 billion a year advertising really necessary?" is that it probably isn't, but that it is pretty much an inevitable consequence of the economic system we live with, which is married to a political system that I like.

Advertising provides one of the clearest lines of visual differentiation between the Communist and non-Communist worlds. We take for granted the constant nattering of commercial messages through the broadcast media; fat and colorful publications in which information and entertainment are wedged as incidentals among compelling exhortations to consume; billboards, car cards, and point-of-purchase displays. And for us, the absence or scarcity of advertising makes for an impression of grey drabness when we enter the urban landscape of Eastern Europe. Advertising is salient to the pictures in our minds as we contrast life under capitalism and life under the new order: Havana 1960 and Havana 1975; Saigon 1974 and Saigon 1976.

This makes it all the more interesting in recent years to note the rise of marketing and advertising as respectable fields of activity in the Communist world. The rehabilitation of advertising arises from several causes:

First, the rise of discretionary income from the post-war low point at which socialist economies were created in Eastern Europe.

Second, the creation of surplus production capacity for many products, even in the Soviet Union, generating a consequent need to develop markets for them.

Third, the erosion of centralized planning doctrine as different state enterprises enter the same areas of production, and rivalries arise.

Finally, the emergence of economic and political pressures to develop export markets. Communist enterprises must compete with Western organizations in non-Communist countries and are, therefore, forced into market research and product promotion. In some countries, foreign imports of luxury goods are permitted, and so is advertising for them.

Needless to say, the Communist world encompasses a variety of systems:

Yugoslavia and Hungary have private economic sectors of modest size in which there is competition, and individual promotions are permissible. Even in comparatively monolithic societies, like East Germany and Bulgaria and the Soviet Union, brand name identity is beginning to assert itself. In Rumania, different factories manufacturing similar merchandise lines have their own retail outlets and vigorously compete for the consumer market, rewarding their managements and employees with appropriate material incentives.

Yugoslavia, with its workers' self-management system of enterprise, has brand name competition and promotion at a level close to that in the West, and not surprisingly, advertising has developed more rapidly there than anywhere in the Soviet bloc. At the same time, as I discovered at an international conference held

last Spring under the auspices of the Yugoslav Advertising Association, there is a strange schizophrenia between practice and theory. The papers and presentations at this meeting included very practical discussions of market planning, advertising department operations, and showings of ads and commercials in no way different from ours. These alternated with a series of "theoretical" papers by Communist functionaries who told the assembled admen that advertising must avoid petty bourgeois tendencies and advance the interest of the toiling masses.

One such paper was delivered by a gentleman who was described to me as a former official in his country's secret police and who now serves as dean of the School of Journalism at a major university. His thesis was one with which many FTC staff members would be in agreement: Advertising serves a useful economic function if it is informative for the consumer. Advertising that appeals to reason and facilitates more rational consumer choice helps promote the sale of desirable goods and services. Advertising is harmful when it appeals to emotion and plays on unconscious motives to persuade people to buy things they do not need.

Following each of these sermons, the usual collection of commercials was shown, with jingles chanting the merits of apparently identical brands of shampoo, canned peas, etc. At a reception later, I asked my new friend, the "red dean," whether he agreed that economic productivity was increased by competition among enterprises producing similar products, like shampoo and canned peas. He agreed that this was so. Was it, then, economically desirable for a firm to try to make its brand or company name familiar to the consuming public? Yes. Could this kind of familiarity be a desirable end in itself even if the consumer was given no other information about the product? He thought for a minute and then gave his official answer: "Unofficially, yes."

Once we accept the principle of competition among firms producing what is essentially the same commodity, it follows that advertising--perhaps not all, but certainly a good deal of it in that product class--will be non-informational, except for the very fundamental information that the brand exists. When brand image represents the only distinctive feature the advertiser has to sell, it is inevitably more likely that he will use irrelevant and non-rational appeals and rely on gimmick techniques to capture attention and heighten identity. In practice, the informational and non-rational aspects of advertising content are very difficult to separate. To the extent that socialist societies rediscover the economic incentive value of competition among enterprises, their tolerance for non-informative advertising content must grow more similar to ours.

This paper addresses the issue of whether that tolerance is justified. I shall try (1) to distinguish between incidental and fundamental questions about non-informational advertising, (2) question whether it is possible to make generalizations out of which regulatory guidelines can be constructed, and (3) point to what I consider a basic error in a widely prevalent assumption that non-informative advertising works in proportion to its reach and frequency.

But how does the traditional Marxist critique differ from that of John Galbraith, Nicholas Kaldor, and other non-Marxist economists who are critical of advertising and of the consumer-oriented economy?

The Marxists regard the failings of advertising as merely incidental manifestations of the evils of capitalism. Advertising is just one of innumerable techniques by which the toiling masses are exploited by the monopoly interests. The critics within our own ranks are also concerned with the exploitative and manipulative aspects of advertising, but they want to "make the system work," as the expression goes, to facilitate competition rather than to eliminate it altogether.

The most commonly heard kinds of complaints about advertising are, in my opinion, incidental to the basic structural criticisms that have been offered in economic terms.

First of all, as Bauer and Greyser's A.A.A.A. study showed a decade ago, objections to advertising very often mask objections to the advertised product. To meet these objections requires limitations on production, distribution, or pricing of a legally available product, or in some instances restricting its use of media with a large audience of children.

Second come objections to specific ads that are esthetically offensive or vulgar and to those whose claims or presentations are false, misleading, or unfairly competitive. Another set of concerns relates to advertising that might be unpersuasive to most mature adults, but that can coerce vulnerable sectors of the population--children or persons of low intelligence or education.

Regardless of how much merit there is in these social criticisms, I don't believe that they concern the basic function of advertising. Both moral law and Federal Trade Commission regulations proclaim that, "Thou shalt not steal, lie, or coerce." Transgressors must be brought to account, but their individual derelictions have very little to do with the whole system of advertising in which we simply cannot assume that theft, deception, and coercion are the norm rather than the exception.

All of these criticisms, I believe, are case-specific and therefore tell us nothing about the essential character of advertising institutions as such.

A third and more troublesome set of objections relates to the influence of advertisers over media content. Apart from the recurring instances in which media dependent on advertising distort their non-advertising content and present publicity and puffery in the guise of disinterested information, it is charged that media censor themselves to avoid offending advertisers and thus do disservice to the consuming public and the body politic. A more subtle and compelling argument is that the array of media content available to the public represents advertisers' choices--or choices made on their behalf--rather than public choices, with disastrous consequences for both public opinion and popular taste.

If we were to accept the argument that the dependence of media upon advertising support has socially undesirable consequences (and the argument must begin by asking "compared to what?") does this represent a fundamental flaw in advertising? I would have to conclude that it does not, since in democratic societies, media are structured and funded in a variety of different ways. As a result, there are substantial variations in the mix of advertising among media, even where the total pressure of advertising is at a high level. In

Germany, for example, television advertising has been deliberately kept to a minimum and in the United Kingdom, radio advertising has only recently been introduced. The British subsidize the BBC through license fees; we subsidize magazines through our low second-class postal rates. Sweden taxes advertising and provides a newsprint subsidy for smaller newspapers. In some countries, the reader pays most of the cost of the newspapers he buys. Here the advertiser does. In this country, advertising of certain products is restricted from certain media by custom, voluntary choice, or law. Broadcasting systems are managed with varying emphasis on maximizing audience size as a prerequisite to gain advertising revenues. All this suggests that while changes and reforms might be made to change the dependence of media on advertising income, these changes would have more to do with the evolution of social policy toward the media than with the functions of advertising itself.

All of the objections to advertising that I have mentioned so far might apply to advertising of both the informative and non-rational kinds and are not related to advertising's economic function.

In evaluating the economic effects of advertising, we may take the standpoint of the individual consumer, of the brand or the firm which manufactures it, of the product class or industry, or of the entire economy. In classical economic theory, all these four sets of benefits interact. That is, if we want to go with Adam Smith rather than with Karl Marx, a consumer benefit is translated into greater consumption, at greater profit for the successful firm, increased sales volume for the product category, and greater efficiency for the whole system through economies of scale in production and distribution. Product innovation is stimulated by advertising, creating new markets and expanding old ones, with consequent benefits for the consumer.

To see where the critics disagree, let me quickly review their two principal themes: You will note that both focus on parity products characterized by high advertising-to-sales ratios.

The archetypal illustration of a parity product of this kind is a campaign for blended whiskey that ran some years ago. A friend of mine who was the copywriter on the account was simply unable to find any good copy points to distinguish his client's product from any of its competitors. Having sought in vain for evidence of superiority, he came up with a slogan that the client loved and that ran for years. It was powerful without being in the least deceptive. The slogan was, "If you can find a better whiskey, buy it."

The first criticism of this kind of advertising for the sake of name registration is that it is economically wasteful and unproductive, that it raises the cost of goods to the customer without providing any commensurate tangible benefit in return. In economic terms, it results in a misallocation of resources.

Nicholas Kaldor argues that the economy spends too much on advertising because it is supplied jointly with goods and services and therefore consumers are forced to buy more of it than they want.

Translated from economic to social terms, the argument is that advertising corrupts and debases values, it creates a preoccupation with material goods and exploits irrational and neurotic motives to promote products that serve little purpose or are actually harmful. Advertising changes social priorities by encouraging expenditures for individually consumed articles and services at the expense of public goods that must

be consumed in common. It persuades consumers to buy things they don't need instead of devoting a larger part of their productive labor and income to the social purposes that everyone needs: clean air and water, crime prevention, the arts, efficient public transportation, preventive medicine, and the like.

A second argument is that advertising contributes to the growth of monopoly and thus to economic inequities. Advertising enhances the power of the dominant firms in product fields characterized by minute differences among brands and in which advertisers compete in terms of brand image rather than through product improvement.

The advertiser of a parity product must promote image rather than information about product advantages. Thus, the sheer weight of his advertising, rather than its content, becomes a prime determinant of his market position. The leaders can set up the biggest budgets and hire the outstanding talent to manage their schedules and create their ads. As they swing their weight around, they can buy advertising most efficiently, outspending and out-maneuvering their competition, progressively increasing their market share. Essentially, this is the conclusion reached by former Assistant Attorney General Donald Turner. When he was head of the Anti-Trust Division, he argued that broadcast television discount structures were a strong inducement to corporate mergers and acquisitions in package goods fields where advertising represents a substantial proportion of corporate expense. It is also the conclusion reached by William Comanor and Thomas Wilson, who, in a recently published econometric analysis, conclude that a firm's ability to dominate advertising in its product category leads to increasing market share, monopoly profits, and barriers to the entry of new competitors.

(It must be noted in passing that while barriers to entry may faze the small entrepreneur, they hardly matter to a conglomerate ready to move into a new field of activity.)

Henry Grabowski and Philip Nelson have taken issue with Comanor and Wilson's data base, analytical procedure, and conclusions, and I have been told that the authors are re-examining their argument. Grabowski concludes that sales explain advertising rather than the other way around. He reports that "with the exception of a few advertising-intensive categories, advertising had an insignificant effect on consumer demand" for "fairly broad aggregate categories." In short, this denies the thesis that it changes consumers' priorities, even though those few advertising-intensive categories may be precisely the ones that attract the lightning bolts. "The few categories in which advertising continued to exhibit highly significant effects on demand ... were characterized by very high average advertising intensities." ... "This suggests that certain types of goods may have product characteristics which are uniquely amenable to stronger advertising effects."

What are these types of goods? Nelson's paper seems to suggest the answer: The lower the utility value of a product, the more heavily it will be advertised and the more likely it is to become a repeat purchase. This analysis "assumes that consumers are able to determine the utility of a brand after enough purchases of that brand."

Yale Brozen and others have argued that "while advertising can persuade a consumer to make the first purchase, no amount of advertising is going to sell the product a second time if the product doesn't deliver." But this argument side-steps the point, which is not that advertising makes unacceptable products look good,

but that it makes acceptable products that are heavily promoted seem preferable to equally acceptable products that aren't--and for no good reason.

In how many product fields does the consumer's choice reflect an assumption that brand differences are real and important, rather than an assumption that all familiar brands are acceptable? In the latter situation, can we assume that the informational content of the message is totally meaningless, and that the only residual value of the advertising is the reminder that the brand exists? To the degree that this is true, advertising is a means of registering the brand name rather than a means of competitive persuasion. This suggests that nonsense is accepted by the consumer as nonsense, with no particular resentment. If this is so, it would appear to support the proposition that the sheer volume of advertising pressure in dollars is directly translated into sales. That proposition is untenable.

The evidence from a number of studies I have made suggests that brand preferences are distributed identically among people who have little or no exposure and those who have a great deal of exposure to the advertising for mass-market parity products. Is the explanation that even the supposedly "unexposed" have some fleeting contacts with those brands, and that these contacts occur randomly in the same proportions as among the more heavily exposed, with each brand's advertising cancelling the effects of each other's? Advertising exposure for any given brand is an index of its visibility in the marketplace. This reflects its sheer physical distribution, which in turn is a function of its market share.

(1) If heavy advertising in a product class could successfully change consumption patterns to its advantage, one might expect that this would lead to increased sales for the category, and hence to a larger share of all advertising.

(2) The fact is that parity products' share of all advertising has dropped by one-fourth in the last ten years, from 29% to 22%. In 1964, groceries, toiletries, toilet goods, beer, and cigarettes represented 47.0% of all national advertising in the measurable media (TV, magazines, and newspapers). Last year, they represented 39.8%. (Incidentally, cigarettes account for only 1.6% of the 7.0% drop.) This decline took place in spite of the fact that these categories have been heavy users of television. Television increased its share of all national advertising from 22% to 25% in the same period, and probably doubled its output of commercial messages, though individual viewing continued at the same level.

(3) Moreover, national advertising has grown at a slower rate than the local sector. In the last ten years, 1964-74, national fell from 61.6% to 55.1% of the total.

(4) This is especially interesting because critics of the non-informational aspects of advertising are generally willing to exclude retail and newspaper classified advertising from their complaints, since it is generally highly specific and factual.

(5) In 1964, national advertising was equivalent to 3.33% of all retail sales volume. Ten years later, the figure was 2.74%--or an 18% drop in the ratio.

Over a ten-year period between the last two reports from the Internal Revenue Service, advertising-to-sales ratios for all manufacturing companies fell from 1.42% to 1.36%, and this decline is visible in most industries in which one expects to find a high proportion of

parity products. Of eight categories of food and beverages, sugar and grain mill products were the only two in which advertising did not decrease as a percent of sales. In tobacco, prior to the broadcasting ban, advertising-to-sales went down. In drugs, it went down. In fact, the few categories where the ratio went up--books, footwear, watches and clocks, and radio and television sets--are all characterized by great fractionation and diversity, and a correspondingly high level of informational content in their advertising.

By the way, have these trends had any visible effect on industry concentration ratios? Between 1963 and 1970, advertising/sales ratios fell from 5.0% to 3.8% for the major cigarette companies and from 13.2% to 10.9% for the leading soap companies. Yet in that period (before the ban on their broadcast advertising and during a period of extraordinary brand volatility) the four leading cigarette companies increased their share from 80% to 84%, while the four leading soap companies' share fell from 72% to 70%. So the monopoly thesis is not easily demonstrable.

But the declining advertising-to-sales ratios raise an even more interesting question, when we look at the efficiency of advertising from the standpoint of the whole economic system. Are advertisers concluding that the ever-increasing number of messages have taken us up against the perceptual threshold beyond which consumers can no longer absorb them? Does this, then, mean that they will increasingly turn to alternative means of promotion as a substitute for advertising? Economic law would suggest that such alternatives will be tested and pursued if their payoff is greater.

It is sometimes suggested that the total amount of advertising in a particular product field should be limited at some arbitrary cutoff point, defined either at a percentage of gross sales volume or at a percentage set in relation to advertising for all products and brands.

A few years ago, in an effort to demonstrate that advertising represented an unwarranted and uneconomic exploitation of the consumer for advertisers' private benefit, the British government forced each of the two major soap companies to stop advertising two leading brands of detergent and to lower the price by the amount of money saved. These brands quickly lost market position, in spite of their 20% lower price.

But limiting advertising for a few particular brands is very different from limiting advertising for a whole category. (In a way, that has been attempted in this country, when utility companies have been restricted from promoting increased consumption of electricity at a time of energy shortages. But this is a special case.) The aggregate promotional effect of all advertising in a product category is competitive with that of all other categories for its share of total consumer spending. Any government body that determined that people should spend less money on sleeping pills, soap, or cosmetics could limit the market much more effectively through discriminatory taxation than by regulating any one particular form of sales promotion.

Limitations on the volume of advertising could simply divert sales pressures into other forms of merchandise and promotion--from market research to push money--directed to the same purpose and perhaps more difficult to regulate. The cigarette industry's diversion of funds out of broadcast advertising into premium offers and auto races represents just one example. Advertising and other promotional activity are hard to disentangle. Cooperative advertising and coupon advertising are often paid out of promotion rather than

advertising budgets.

Regulations and controls of advertising have to be based on generalizations. That is, you have to be able to assume that you're dealing with a class of events in order to subject such events to regulatory proceedings. Frank M. Bass (1974) has recently pointed to the difficulty of generalizing about advertising performance. He reanalyzed a Federal Trade Commission study of 97 food companies and found that in only one-third of the cases there was a strong link between profitability and the advertising-to-sales ratio. Moreover, he writes, "There are some industries in which advertising and profitability are independent, while there are other industries in which advertising and profitability are strongly related... the differentiating characteristic ... is not the size of the advertising/sales ratio."

Comanor and Wilson also come up with qualified answers when they review the evidence of advertising's effect on the market share of the leading brand. They find considerable variations from one product class to the next.

As a matter of fact, my own look at different product categories shows that the percentage of national advertising accounted for by the leading company or companies varies all over the lot. It may be necessary for a firm to register a minimum amount of visibility or acceptability through advertising pressure. Still, it is difficult to state with any precision at what point such a barrier to entry is transformed from a minor obstacle to an impenetrable wall.

Why is it so hard to generalize about advertising? Because its unpredictable creative element cannot be programmed. Looking at advertising volume strictly in terms of dollar investments gives us no real clue as to its yield. We're back with John Wanamaker's old observation that 50% of his advertising was wasted, although he didn't know which 50% it was. My own guess is that if John's experience was typical, 50% was an extremely low estimate.

In the economic studies, it is tacitly assumed by both the critics and supporters of advertising that creative quality may be considered a constant; that its variability shows about the same range that you would expect to find in product performance or in entrepreneurial talent. This assumption is critical, and also mistaken.

The striking variability in advertising productivity must lead us to conclude that the yield from advertising is not actually related to its sheer volume of expenditure or to its ratio to sales. The Universal Product Code and the computerized supermarket checkout have made it possible to determine precise contributions of advertising to product movement over a very short period of time. We have been tracking weekly sales, item by item, in a number of supermarkets with automated checkout counters, in different parts of the country. In one Ohio chain, we compared unit sales for the week in which an item was listed in a retail food ad with its sales during other weeks of the month in which it was not advertised. This represents a true measure of the effects of only two elements: (1) the amount of advertising pressure and (2) the attractiveness and value of the product to the consumer.

One private label brand showed a 614% increase for the test week in which it was advertised. Another product, advertised with exactly the same weight and prominence, showed a 22% increase in sales for the test week. (In this particular instance, the difference in demand elasticity for the two products may be the best

explanation of this thirty-fold difference in the rate of sales increase.)

Tracking items that had a much smaller listing in a supermarket ad, we found one that showed a 180% increase and another that actually showed a 25% decline during the advertised week, perhaps because of competitive promotions. Taking two coupon offers of identical size and typography buried within the same ad, one was followed by a 782% sales increase, another by an increase of 45%.

Coupons for a typical grocery item involves a very small purchase decision. But the U.S. Army recently (1974) used an inquiry coupon in recruitment advertising bound into forty-four magazines. The cost per inquiry ranged from \$4.68 to \$393.28, and the cost per resulting enlistment ranged from \$165.42 to \$84,050.

Apart from variations in the suitability of media for specific purposes, different advertising messages for the same product vary tremendously in persuasive appeal and power, as copy testers have long known. Evidence continues to accumulate that many advertising messages are forgotten very rapidly. For example, a study recently made by W. R. Simmons for Golden West Broadcasting shows that among telephone respondents who had been watching a TV station for the previous hour, 1.5% of all the commercials in that hour were recalled (with aid).

No doubt newspapers and other advertising media have an equivalent rate of forgetfulness for messages that are essentially delivered at random, rather than actively sought out.

If only a small proportion of ad messages are remembered, this implies that an incredible amount of waste must occur in advertising communication. I do not mean to suggest that a message that is not consciously remembered is without effect, since its echoes may be evoked when it is repeated or when the consumer confronts the product at the point of sale. But messages that are poor in memorability or inadequate in persuasive content are more likely to be non-productive, under-productive, or counter-productive for the advertiser and for his product field, and thus serve no useful economic function. What can be the social benefits of this unproductive advertising, which fails to produce sales effects commensurate to its cost, and therefore results in no increased economic activity?

So--Is all this advertising really necessary? Well, yes--and no.

(1) It is impossible to conceive of a competitive market economy in which advertising does not play a part. Advertising is one of a variety of forms of selling and promotion. The advantages of size to an advertiser parallel advantages of size in other realms of merchandising, distribution, and sales. There is no logical or reasonable way of restricting one aspect of promotion without restricting all other ways--which would mean the end of competition.

(2) In many instances and in many aspects, there may be legitimate reason to criticize the economic and social effects of advertising. But these instances must be dealt with on their own specifics. A knee-jerk defense reaction to any attack on any aspect of advertising is just as witless as knee-jerk hostility. The economic effects of advertising are not subject to universal generalizations.

(3) A growing percentage of all advertising in informational in character. A declining proportion is

represented by national advertising in product fields characterized by minor performance differences among brands. This trend should continue as we have increasing product differentiation in the market place and more rapid growth in the consumption of services than of goods.

(4) The misallocation of resources that results from overspending by certain advertisers is probably small relative to the misallocation that arises from inadequacies of advertising content and scheduling. To correct those misallocations is precisely the task of consumer research.

#### REFERENCE

Frank M. Bass, "Profit and the A/S Ratio," Journal of Advertising Research, 14(December 1974), 9-19.

## INTERPRETING THE POLLS

Seymour Martin Lipset, Stanford University

Public opinion research is in both good shape and bad shape. On one hand, it is probably more generally accepted and used by various buyers than ever before. Business, governments, politicians, labor, all now commission surveys before they start a "sales campaign," or to evaluate how they are doing with segments of the public. They want a breakdown of the public's values and interests with respect to what they are selling or doing. The survey has become a major tool of academic research. Government agencies now use samples in lieu of population censuses. Our breathlessly awaited labor force figures are the product of sample surveys.

Yet there are also problems. Just the other day the press carried a story about a situation increasingly familiar to us, the reluctance of the public to be interviewed. The refusal rate has increased. It costs more to find respondents. The same pattern is evident for questionnaire surveys. In discussing the phenomenon, some researchers have raised the question as to whether or not this increased resistance reflects growing public hostility to polling, an increase in the sentiment that it is no one's business what I think, perhaps a linkage of legitimate polling to its abuse by government security agencies or sales firms who have had their agents pose as interviewers to gain access to homes or to secure confidential information. The increased concern about privacy has led some to query whether a polling organization can in fact guarantee anonymity. Some time ago, I had an argument with Vern Countryman of the Harvard Law School who had publicly urged students not to complete questionnaires sent out by the Carnegie Commission on Higher Education on the grounds that any information they supplied concerning their political views or affiliations or their use of marijuana or other illicit drugs might end up in the hands of police agencies. When I challenged Countryman's assumptions, saying this had never happened, that pollsters like physicians, lawyers, or newspapermen would refuse to divulge sources, he contended there was nothing in the law to protect their right to do so, and that he so advised everyone who was asked for an interview or to fill out a questionnaire.

Fortunately, for the survey business, this legal view has not penetrated widely, but still one wonders as to whether or not the combination of the increased suspicion and awareness of governmental and business invasions of personal privacy, combined with the use of the interviewer role as a means of gaining access by salesmen, in the context of an increasingly suspicious world, is responsible for the growing refusal rate. (This reminds me of one definition of the difference between the paranoid and a normal person. The paranoid is a person who lacks the normal person's ability to repress awareness of the hostile behavior of those around him. Perhaps we are all becoming more aware.)

But if the public in the potential role of interviewee is less willing to cooperate, the public in the role of consumer of opinion research is eager to learn. As noted before, various elites very much want to know how the people out there are reacting. The growing numbers of college educated among the population are also desirous of knowing. And as elites have become more concerned with opinion findings, opinion researchers have increased their influence on outcomes. This is not the old issue of the band wagon effect among the

electorate, one which has never been resolved, a fact which should mean that the effect in any case is not very great. But where opinion results clearly do matter is in affecting the behavior of decision-makers who change their policies or their strategies because of survey findings.

We can see this today in the current presidential campaign. The published polls have almost made a front-runner out of the most eager non-candidate in history, Hubert Humphrey. Associates of the announced Democratic candidates complain bitterly over the fact that Gallup and Harris run presidential trial heats which include Kennedy and Humphrey on the list of possible candidates. They say that this is unfair, that the pollsters should take these men at their word, that they should limit the choices to those who are overtly in the race. In effect, they are arguing that Gallup and Harris have created a situation which bypasses the normal primary system, that the polls may determine who the Democratic nominee will be, that if published polls were not available the situation would be different.

The new conservative strategy advanced in recent books by William Rusher and Kevin Phillips urging the creation of a Conservative Party to replace the Republicans, which underlies some of the enthusiasm behind the campaign for Ronald Reagan, is in large part a reaction to the surveys which show that there are many more self-identified "conservatives" than "liberals" at the same time that self-proclaimed Democrats considerably outnumber Republicans.

Poll results in pre-primary situations, which are inherently much less reliable than in general elections, because of less knowledge and structure about the choices, and the small proportion of the electorate who vote in primaries, have helped to determine the results of the primaries by undercutting the funds and enthusiasm for candidates who appear as also-rans in these surveys. Hubert Humphrey was greatly disadvantaged in the decisive California primary in 1972 by published surveys which showed him further behind McGovern than the actual results suggested he had been. (Parenthetically, I would note that there have been much greater discrepancies between primary election returns and poll estimates in a number of situations, a fact rarely if ever mentioned in newspaper columns presenting pre-primary public preferences. Discrepancies between the anticipations of pre-election surveys and the actual election returns are lowest in Presidential contests where the pollsters report on the opinions of voters at the end of a highly publicized well structured partisan contest which has lasted for months and in which turnout is higher than in other situations. Efforts to "predict" the results of primary contests or of non-partisan mayoralty races have produced gaps of from 10 to 20 percent. In Minneapolis, for example, the Metro Poll published by the Star reported mayoralty candidate Charles Stenvig running third in the elimination primary in 1969. Stenvig actually placed first by a wide margin. In 1973, the Star's poll considerably over-estimated Stenvig's support, seeing him in the lead before the elimination by 10 percent when he lost the race. In Boston, this year, the Globe's mayoralty poll for the preliminary September primary, published two days before the election, reported incumbent Kevin White ahead by 2:1, with a lead of over 20 percent.



White's actual plurality was less than half of this. The Globe decided not to run a poll for the final election. In the November, 1975 elimination primary for mayor of San Francisco, candidate Dianne Feinstein's strategy "to conduct the high tone campaign aimed at issues, not personalities" was, according to the San Francisco Chronicle, "based on polls which showed Mrs. Feinstein comfortably assured of a spot in the run-off." In the election she finished in third place out of the running for the final contest.

In a different context, it may be argued that the sense of malaise in the country, that we are experiencing a "failure of nerve," a crisis of legitimacy or authority, is heightened by the survey data, indicating decreasing confidence in various institutions, the Presidency, the Congress, business, religion, etc. These results encourage the left radical critics much as the preponderance of "conservatives" encourages the right-wingers.

Opinion polls have power to affect what happens. This should make us especially concerned about issues of validity and reliability, and at least equally important about the way results are presented, how they are interpreted. There is, of course, no way of controlling how those who read the results of a survey interpret them, but opinion researchers do have an obligation to point out the way in which different factors may affect the reliability of their findings, e.g., sampling issues, non-respondents, wording of questions, choices given to respondents, the place a question is located, the saliency of issues, the lack of stability of various opinions, etc.

These issues occasioned some public comment recently when the press noticed that Gallup and Harris produced somewhat different estimates of the public's evaluation of President Ford's performance. In a survey released for publication in early December 1974, Gallup reported that in answer to the question: "Do you approve or disapprove of the way Ford is handling his job as President?" 42 percent approved, 41 percent disapproved and 17 percent had no opinion. Harris' December results for the question: "How do you rate the job President Ford is doing as President--excellent, pretty good, only fair, or poor" found 46 percent giving "positive" replies, 52 percent negative and only 2 percent "not sure." The differences between the two surveys upset some newspaper editors, but it should be noted that such differences had occurred on a number of previous occasions. Thus, Gallup's November 1974 figures for the same question were 48, 32, 20, while Harris' for his formulation were 48, 47, 5. For October, Gallup's reported a large approval majority, 55, 28, 17, while Harris found more disapprovers, 45, 49, 6. This latter difference is quite startling. Gallup indicated the country approved of Ford by 55 percent to 28, while Harris said he was disapproved by 49 to 45. The variations between the two surveys in reporting the percentage who had no opinion are consistently of a considerable magnitude, Gallup's running from 15 to more than 20 percent, while Harris generally come to well under 10 percent.

Variations such as these occurred with respect to estimates of the popularity of Congress as well. In March 1975, Gallup reported 32 percent approved and 50 percent disapproved "of the way Congress is handling its job." Harris indicated that 26 percent had a positive news of "the job Congress has been doing so far this year," while 67 percent were negative, that is negative views outnumbered positive ones in Gallup's data by 18 percent, while according to Harris' results, the difference was 41 percent.

If we look at the public's view of the way the President has handled specific issues, comparable differences were reported by the two newspaper surveys. A Harris survey taken in December 1974 found 60 percent giving President Ford a negative rating for his economic program, as compared to 48 percent negative in a Gallup survey.

The issue of discrepancies between polls has again become a matter of considerable importance when the New York Times found in a national survey which it conducted at the beginning of November 1975 that the public, by 55 to 33 percent, favored federal funds to help New York City in its financial crisis. Three weeks earlier, before President Ford's speech opposing such a policy, Gallup had reported that the opponents dominated by 49 to 42 percent. The variation over a three week period between a seven percent plurality disapproving and a 22 percent approval, one occurred in response to answers to the identically worded question: "Do you think the federal Government should or should not provide funds to help New York City get out of its financial difficulties?" The two surveys, however, employed different interviewing techniques. Gallup's results were obtained from face to face interviews, while the Times queried its respondents by telephone.

Such differences with respect to specific issue questions, have showed up repeatedly, particularly when the question is worded somewhat differently. Thus, in January 1975, Gallup asked respondents to choose between "two plans to reduce consumption of gasoline," Plan A involved rationing, "Each driver would be able to buy up to 10 gallons per week with the price remaining at the amount he or she presently pays;" Plan B contained a price increase, "Each driver would be able to buy as much gas as desired, but the price would be increased by about 10 cents per gallon above what he or she presently pays." Harris, at about the same time asked: "In order to conserve oil, if you had to choose would you have mandatory gasoline rationing, on an odd-even basis with no increase in the price of gasoline, or no rationing, but an 11-cut-a gallon rise in the price of gasoline and fuel oil as a result of the tariff on imported oil from abroad?" Both reported their results under the headings, rationing compared to a price increase or oil imports tax. Gallup found 37 percent for rationing, 48 percent for a 10 cent per gallon price increase, while Harris reported 60 percent for rationing and only 25 percent for an 11 cent per gallon increase in the tax. Much of the difference, of course, probably was the result of the variation in question wording, one defined rationing as limiting purchases to 10 gallons a week, while the other did not mention any limit. But the reports of each as published in the press presented the results in terms of support for rationing versus a price increase, and in these terms, they reported sharply contradictory results.

Again, in December 1974, Gallup asked whether people expect 1975 to be a year of "rising prices or... of falling prices." Harris inquired as to whether "a year from now" you expect prices to be rising more rapidly, less rapidly, staying the same or going down. Gallup found 19 percent expected falling prices, Harris' figure was 4 percent.

The results of the ideological self-identification question which has so impressed William Rusher and Kevin Phillips also has been responded to differently in surveys conducted about the same time. Between November 1974 and March 1975, four different polling agencies asked national samples to locate themselves as conservatives or liberals, asking differently worded questions. Each reported varying distributions.

Harris and Gallup found many more self-identified conservatives and liberals while the two university polling centers, the Survey Research Center (S.R.C.) of the University of Michigan and the National Opinion Research Center (N.O.R.C.) of the University of Chicago reported little or no differences in the proportions of the population identifying with each term. In November, Gallup asked: "if an arrangement of this kind, that is two new political parties were carried out, and you had to make a choice, which party would you personally prefer--the conservative party or the liberal party?" The results were 40 percent conservative, 30 percent liberal, and 30 percent undecided. In December, Harris inquired: "How would you describe your own political philosophy, as conservative, middle-of-the-road, liberal or radical?" His findings were 30 percent conservative, 15 percent liberal, 3 percent radical, 43 percent middle-of-the-road and 9 percent not sure. In November also, S.R.C. interviewers told respondents: "We hear a lot of talk these days about liberals and conservatives. I'm going to show you a seven point scale on which the political views that people might hold are arranged from extremely liberal to extremely conservative, where would you place yourself on this scale or haven't you thought much about it?" Conservatives led liberals slightly by 25 to 21 percent, with 26 percent choosing the mid-point position, 21 saying they haven't thought much about it, and 7 giving don't know as their response. The question asked in March 1975 by N.O.R.C. was almost identical to S.R.C.'s with the exception of the fact that they did not include the option, "or haven't you thought much about it." Out of this format their study produced a tie between the liberal and conservative alternatives. Both received 28 percent, with 37 percent placing themselves in the middle and 6 percent responding don't know.

Lest it appear that I am picking on the commercial polls, I should note that one can find additional variations in reports on the same issues in the surveys of the two main academic organizations. In their 1972 omnibus survey, the N.O.R.C. found 20 percent favored "busing of Negro and white children from one district to another," while 77 percent were opposed. S.R.C.'s 1972 survey indicated that only 9 percent supported "busing for integration" as compared to 86 percent for keeping children in neighborhood schools.

One of the organizations, Michigan's Survey Research Center, furnished an interesting set of sharply divergent results in two of their own studies, taken it should be noted two months apart with the 1968 Presidential election intervening. In the pre-election survey, the Michigan interviewers asked a national sample: "Would you say that people like you have quite a lot of say about what the government does or that you don't have much say at all?" Three quarters chose the option, "don't have much say at all." Two months later, the same issue was presented as an agree-disagree item in the following terms: "People like me don't have any say what the government does." This time only 41 percent agreed or 34 percent less than seemingly had voiced a comparable opinion earlier. A difference of the same magnitude in replies to highly similar questions in the two S.R.C. polls was reported in answer to the questions a) pre-election--"Would you say that politics and government are so complicated that people like you can't really understand what's going on or that you can understand what's going on pretty well," and b) post-election, "Sometimes politics and government seem so complicated that a person like me can't really understand what's going on." Less than half the respondents, 44.5 percent, chose the alternative, "can't really understand" in the first survey, while 71 percent agreed with this view two months later.

Some might suggest that the reason for the varying responses is that many less informed or uncommitted persons are inclined to agree with a statement, and in both cases the change in views resulted from changing the question from an either/or form to agree-disagree. This may be so, but the same surveys also included two other closely related questions which shifted format, but not response distribution. Thus, the S.R.C. pre-election study asked, "Would you say that most public officials care quite a lot about what people like you think, or that they don't care much at all?" Following the election S.R.C. included the item: "I don't think public officials care much what people like and think." The two formulations produced highly comparable answer patterns. Forty percent chose the "don't care" option in the first study, while only four percent more agreed that "public officials don't care" in the second. And to reiterate the finding and to demonstrate that the sharp variations to the first two sets of questions did not result from disgruntled supporters of the defeated candidates, Hubert Humphrey and George Wallace, changing their minds, it may be noted that questions about voting in the two studies produced almost identical response patterns. Thus, 58 percent chose the option, "voting is the only way that people like you can have a say about the way the government runs things," or an either/or item, while 57 agreed, "voting is the only way" when presented with this view in the agree/disagree format.

The differences in response to these four sets of parallel questions asked two months apart are puzzling to say the least. I have no plausible explanation.

If we turn to the crucial area of foreign policy, the same pattern of response variations reoccurs. In April 1973, Harris found 49 percent opposed the "bombing by U.S. planes in Cambodia." During the same month Gallup's results indicated 57 percent disapproved "bombing Communist positions in Cambodia and Laos." The two most widely published surveys also produced varying results this same year with respect to opinions about the way that the United States should treat North Vietnam after the war. In January, Gallup reported in reply to the question: "if a peace agreement is reached should we help rebuild North Vietnamese cities?" That 42 percent were in favor of so doing. In February, Harris inquired as to whether respondents "favor aid to North Vietnam to rebuild war-time damage?" and found only 21 percent support for such a policy. There was a similar lack of consensus in estimating the public's attitude toward reestablishing diplomatic relations with Cuba at the end of 1974. Gallup reported 63 percent in favor; Harris indicated only 50 percent had that view.

One of the sharpest variations in sentiments on an important policy issue was reported over a five month period by the Harris organization. In December 1974, Harris interviewers asked: "There has been a lot of discussion about what circumstances might justify U.S. military involvement, including the use of U.S. troops. Do you feel if (12 different circumstances for various countries) you would favor or oppose U.S. military involvement?" Only 14 percent favored such involvement "if North Korea attacked South Korea," while 65 percent were opposed. In May 1975, Harris asked a specific question about Korea in the following terms: "The U.S. has 36,000 troops and airmen in South Korea. If North Korea invaded South Korea, we have a firm commitment to defend South Korea with our own military forces. If South Korea were invaded by North Korea, would you favor or oppose the U.S. using troops, air power and naval power to defend South Korea?" Not surprisingly, this wording elicited a much higher positive response for American military participation

than the earlier question, but still the variation was staggering, 43-37 percent were for the use of U.S. troops in the second study, as compared to the earlier unfavorable majority of 65-14 percent.

Reports on public opinion toward U.S. involvement in Korea were equally disparate twenty-five years ago during the Korean war. Thus, in December 1950, the Gallup Poll found in answer to the question: "Do you think the United States made a mistake in going in the war in Korea or not" that only 39 percent approved of the war. An N.O.R.C. survey taken in the same month reported 55 percent in favor in response to the query: "Do you think the U.S. was right or wrong in sending American troops to stop the Communist invasion of South Korea?" In April 1951, a repeat by Gallup of his earlier question produced a 43 percent vote in favor, while N.O.R.C. found 63 percent for our intervention in reply to their more emotionally worded question. Two surveys conducted by N.O.R.C. one month apart in 1953 produced a 37 variation in the proportion supporting the war, seemingly as a result of varying the question. In August, they asked: "As things stand now, do you feel the war in Korea has been worth fighting, or not?" Only 27 percent responded positively. But one month later, when N.R.C. used the earlier formulation as to whether "the United States was right or wrong" to have sent troops, 64 percent said the policy was right. Can one speak of the true or real feelings of the American public with respect to Korean policy in the fifties or now when the same polling organizations, presumably using the same sampling frames, interviewers, etc. can elicit such disparate results by varying the wording of the question.

The ability of pollsters to change response patterns and to find sharply different sentiments on what appears to be the same issue can also be illustrated with respect to attitudes toward American policies to the Middle East conflict. In presenting the results of a survey taken in December 1974, Harris wrote in a New York Times Magazine article: "Another lopsided 66 to 24 majority favors sending Israel what it needs in the way of military hardware." One month earlier a Yankelovich poll found only 31 percent in favor of the U.S. sending arms to Israel, while 57 percent were against. In January 1975, Yankelovich found 45 percent in favor of military aid to Israel in response to one question, a figure which declined to 28 percent when the question was formulated differently in the same survey. But it must be reported that a Gallup survey also taken in January reported only 16 percent supported military aid of various types for the Jewish state with another 7 percent urging general support. Over half the respondents, 55 percent, gave Gallup interviewers responses which were coded under the heading, "stay out of the conflict." A couple of months later, however, Gallup reported that 54 percent favored either sending military supplies (42 percent) or American troops (12 percent), while only 37 percent opposed American aid to Israel in a renewed Middle East conflagration.

These drastic variations seemingly resulted from the very different way the questions were formulated in the five studies. Harris' December interviewers elicited a 66 percent positive response for military aid to Israel when they asked the question: "As you know, the U.S. has sent planes, tanks, artillery, and other weapons to arm Israel. The Russians have sent similar military supplies for Egypt and Syria. In general, with the Russians arming Egypt and Syria, do you think the U.S. is right or wrong to send Israel the military supplies it needs?" Yankelovich found his 31 percent figure in November in reply to a question about military aid to Israel in the context of queries about a

number of countries: "The U.S. sends arms and military equipment to a number of foreign countries. Do you personally feel that the U.S. should or should not send arms to (country A, B, C, Israel)?" His 45 percent favorable to military aid in January was in reply to "In view of the situation in the Middle East, do you feel that U.S. should increase its present aid to Israel, continue it at the same level as now, or cut it back." Thirty-six percent said, "continue," while 9 percent favored an increase. The much lower 28 percent figure in the same survey was in response to, "Do you favor selling arms and military equipment to both Israel and the Arabs, just Israel, just Arabs, or neither." Fourteen percent said, "Both," another 44 percent, "Just Israel," and almost two-thirds, 63 percent, were opposed to selling arms to either.

Gallup's low report of 16 percent was obtained in January in reply to an open-ended question: "What should the U.S. do if a full-scale war breaks out in the Middle East?" His high estimate of 54 percent occurred in April in answer to the query: "In the event, a nation is attacked by Communist-backed forces, there are several things the U.S. can do about it. What action would you want us to take if Israel is attacked --send American troops, or send military supplies but not send American troops, or refuse to get involved?" These six questions produced percentages in favor of sending or selling arms and/or troops to aid Israel of 66, 45, 31, 28, 16 and 54. And as a final note on this issue, it must be reported that a February 1975 Harris survey found the public opposed to "selling military equipment to (all) other nations" by 53-35 percent.

The problem basically is that public attitudes toward a given issue are usually too complex to be summed up by the responses to one or two questions. People can and do hold what appear to be contradictory opinions on the same subject. This point may be illustrated by reference to the responses to a number of questions on the Middle East situation given by a national sample of American professors in a survey conducted in the Spring of 1975 by Everett Ladd and myself. Almost three-quarters, 74 percent, agreed with the statement, "The U.S. should pursue a more neutral and even-handed policy in the Middle East," of whom 31 percent strongly agreed. Slightly more than half agreed that "The U.S. should apply pressure on Israel to give in more to Arab demands." Two-thirds disagreed with the proposal that "If the United Nations were to vote to expel Israel, the U.S. should withdraw from the U.N. in protest."

Yet over two-thirds of the same national sample of college faculty approved the statement: "The U.S. should continue to supply Israel with weapons and military equipment." When asked to choose among four alternative courses of United States action, "If Israel were attacked by Arab countries and threatened with defeat," less than a quarter, 24 percent, recommended that it "take no military action." A larger group, 30 percent, endorsed either sending "air support" (13) or "U.S. troops if necessary" (17). The remaining 45 percent favored sending "military aid but not U.S. personnel." Over three quarters of the faculty respondents agreed that "Israel has a right to keep the city of Jerusalem as its capital, so long as the Israelis respect the religious rights of Christians and Moslems."

Close to two-thirds, 64 percent, however, thought that "The Arabs should be allowed to set up a separate nation of Palestine on the West Bank of the Jordan." But only 13 percent agreed that "Guerilla activities on the part of the Palestinian Arabs are justified because there is no other way for them to bring their grievances to the attention of the world." This reply, however, did not reflect the large majority's

disapproval of violent means for almost two-thirds, 65 percent, disagreed with the statement, "It is wrong for Israel to retaliate against the Arabs whenever Arab guerillas commit an act of terrorism."

Looking at this diverse set of responses, it is apparent that it would have been impossible to sum up the views of this group by a few simple questions designed to locate their general sympathies in the Middle East conflict. Like most Americans, faculty sympathies lie more with Israel (57 percent) than with the Arabs (8 percent), and such sympathies are reflected in support for armed aid for Israel, for its claims on Jerusalem, for its right to retaliate. But the faculty, also by a large majority do not want to see the U.S. involved in Middle East war, and would like to see the tensions reduced. These sentiments give rise to approval of proposals that the U.S. pursue a more "even-handed" policy, that it press Israel to "give in more to Arab demands," that the Arabs be allowed a Palestinian state.

Presenting this example, drawn from my own work, is not designed to suggest that commercial pollsters are naive about the complexities involved in issues such as these or that they have not done comparable research designed to explore issue opinion in depth. The December 1974 Middle East study of the Harris organization, as well as surveys on the same problem by Yankelovich, and various Gallup studies of foreign policy have included a variety of different questions designed to find the parameters of opinion. The problem is not that the research design is unsophisticated but that the published reports or even private ones for clients often simplify the issues, on the assumption that the reader will not understand, or care to know about complexities, that he wants relatively straight-forward and simple answers with respect to attitudes toward federal help for New York, aid to Israel, proportion interested in buying a new car next year, etc.

There are, of course, many other issues involved in interpreting opinion from surveys, than those I have discussed. The picture presented can appear a great deal differently depending on how the answers to the same questions are presented. Thus, a number of polling organizations have inquired as to the confidence the public has in the people in charge of various institutions, medicine, Congress, major companies, the Supreme Court, etc. All the surveys agree that confidence as expressed in responses to these questions has been eroding since the mid-sixties, although again it must be mentioned that the percentages reported giving the same response to the same question for the same institution at about the same time have varied considerably.

One polling organization, the market research division of Procter & Gamble Company, noting the variation in responses to such questions, recently undertook an experiment to see how much expression of confidence in different institutions may be varied by using different terms to describe them. They divided their sample into three groups and asked them whether they "have a great deal of confidence, a moderate amount of confidence, or no confidence in it" for a number of institutions giving each third a different term for the same institution. Some of the results are presented below.

TABLE 1  
LEVEL OF CONFIDENCE IN DIFFERENT AMERICAN INSTITUTIONS  
USING DIFFERENT TERMS

Institutions	Level of Confidence		
	Great Deal	Moderate	No Confidence
Army, Navy and Air Force	63	31	6
The Military	48	45	7
Military Leaders	21	59	20
Established Religion	50	41	9
Ministers and other Religious Leaders	35	56	9
Organized Religion	35	49	16
Television News	31	58	11
Network Television News	25	69	6
Television News Commentators	23	66	11
Colleges	2	52	46
Professors	9	62	29
Business	20	73	7
Business Leaders	18	68	14
Big Business	12	62	26
Organized Labor	21	53	26
Union Leaders	12	63	25
Big Labor	7	66	27
Public Opinion Pollsters	14	68	18
Public Opinion Polls	16	55	29
Election Polls	20	45	35
U.S. Presidency	30	53	17
Executive Branch of Federal Government	18	62	20
Federal Government	16	69	15
Government	20	64	16
Elected Government Officials	5	75	20
Congress	7	74	19
Politicians	2	53	45
Politics	4	65	31

It is clear from looking at these results that a sharply different picture of the level of confidence in different institutions emerges depending on the words which are used to depict them. Fifty percent have a great deal of confidence in established religion but only 35 percent feel the same way about organized religion. Almost two-thirds, 63 percent, are very positive about the "Army, Navy and Air Force," but the high level of confidence declines to 48 percent for the "Military," and drops way down to 21 percent for "Military Leaders." More than a third, 35 percent, express no confidence in "election polls," but only 18 percent have the same negative view of "Public Opinion Pollsters." Twenty-one percent have a great deal of confidence in "Organized Labor," but only 7 percent have the same view of "Big Labor." The responses reported in this table tell us a great deal about the public's sentiments but as important is the fact that they illustrate in detail the instability of such replies, the extent to which it is possible to vary the public's view by changing the way in which institutions are depicted.

It is also important to recognize that one gets a very different image of the confidence level of the country depending on whether a survey organization reports and discusses the proportion of respondents voicing "great confidence," while lumping together those indicating "some confidence" with respondents selecting the "hardly any confidence at all" to represent their point of view. The most widely circulated poll reports published in the press, those conducted by the Harris organization, only list the "great deal of confidence" figure which in recent years has been under 50 percent, often well under that figure, for the leaders of most institutions. And

this fact is generally interpreted in accompanying commentary to mean that Americans lack confidence in the leadership of almost all their key institutions. But if we look at the percent of those who say they have "hardly any confidence," a quite different picture emerges. For it turns out, according to an N.O.R.C. 1975 survey, that for all except the leaders in the political realm (the executive) and organized labor, the proportion indicating a lack of confidence runs from a tenth to a fifth, while even for politics and labor, the proportions are all under 30 percent. Or to put the matter another way, from 71 to 92 percent of respondents indicate "some" or "a great deal of confidence" in various key institutions. Is the confidence glass more full or more empty? The implications of the results are debatable, but Harris in reporting the low proportions voicing "a great deal of confidence" concluded: "In short, there is a leadership vacuum in this country across the board."

A similar point may be made with respect to the interpretation by Gallup of the differences in the public's response between 1971 and 1975 to a question which was cited earlier with respect to Israel. The same question, "In the event a nation is attacked by Communist-backed forces, there are several things the U.S. can do about it. What action would you want to see us take if (a different specific country) is attacked--send American troops, or send military supplies but not send American troops, or refuse to get involved?" was asked about a number of countries. The results were published in The New York Times of May 11, 1975 with the descriptive interpretation that they offered "little evidence" of the much heralded trend toward isolationism among the public. This conclusion is warranted by the fact that there was very little difference in the percentages favoring sending American troops between 1971 and 1975. The average figure dropped only one percent for the seven nations. The picture, however, is quite different with respect to proportions who chose the option that the U.S. should "refuse to get involved." On the average, the non-involvement group increased by 7 percent, with much more substantial changes for some countries, e.g., for West Germany, non-involvement rose from 22 to 33 percent, for Taiwan from 45 to 54, for Turkey, from 37-49. Clearly, the shifts over four years in the proportion saying we should "refuse to get involved," present strong evidence for the conclusion that the Americans have become more isolationist, but Gallup apparently contended that his data refute such a view because the generally small proportions favoring the sending of troops had not increased.

Given the increased influence survey analysts have in affecting policies of businesses, other institutions, journalists, politicians, and the mood of the general public, it is important that the limitations of the instrument be recognized more widely than they are. Much of this is obvious to those who work professionally with survey data, they know their own weaknesses. But like most businessmen, they do not stress the deficiencies of their product to clients. They do not emphasize the complexities involved in analyzing data, or the need often for more expensive research or more detailed complicated write-ups if the client is to understand the state of opinion.

Thus it is obvious, and George Gallup made the point over three decades ago, that it is easy to change responses by presenting a given view in association with positively or negatively valued symbols, goals or persons. By associating a given choice of action with resistance to Communist attack or with the possibility of sending U.S. troops into action, the percent favorable may be varied by as much as 20 percent. Response distribution changed greatly by where a question is

located in a schedule. Three academic authorities on survey research, Milton Rosenberg, Sidney Verba and Philip Converse, who examined many of the surveys dealing with Vietnam during the war in their book, Vietnam and the Silent Majority, after reporting large-scale variation in different polls seeking to measure support or opposition to the war or to specific Vietnam policies concluded (1970: 23-24):

One of the reasons why subtle changes in the wording of a question produce different responses is that many of the people whom the pollster questions do not have very well-formed and deeply held opinions on the matters about which a pollster is asking. They are likely to be responding to a question to which they have not given much or any previous thought. What this means is that the wording of the question makes a big difference in how they reply. It also means that the answers that any individual gives can possibly change from day to day. If an individual has not given serious thought to a question, his answer is likely to be offhand. If the pollster were to come back the next day, a somewhat different answer would be obtained... If a question is asked in which negative symbols are associated with withdrawal from the war, people sound quite "hawkish" in their responses. Thus, people reject "defeat," "Communist take-overs," and "the loss of American credibility." On the other hand, if negative symbols are associated with a prowar position, the American public will sound "dovish." They reject "killings," "continuing the war," and "domestic costs." Turning the matter upside down, we see the same thing. If positive symbols are associated with the war, the American public sounds "hawkish." They support "American prestige," "defense of democracy," and "support for our soldiers in Vietnam." On the other hand, if positive symbols are associated with "dovish" positions, the people sound "dovish." They come out in support of "peace," "worrying about our own problems before we worry about the problems of other people," and "saving American lives."

Thus it is possible, even in the same poll, to have the American public sounding like hawks and doves at the same time. At times, these seeming inconsistencies are due to the fact that the positions are not that inconsistent, but at times they are due simply to the alternative question wordings. Lastly, we should point out that many Americans do, in fact, hold inconsistent views on the war. They favor sets of policies that are not compatible one with another... (Rosenberg, Verba, Converse, 1970: 25).

In seeking to understand the considerable variations in response to questions about federal aid to New York City, between the New York Times and Gallup results, cited earlier, Robert Reinhold, a New York Times reporter, noted Gallup and the Times used different interviewing methods, face to face interview and telephone, that "the questions were asked in different context," "measurement error"--imperfections in the questionnaire or slight differences in the way questions are asked and perceived, "refusals, 30 percent of those called by the Times "declined to be interviewed," and "wide public confusion over the complex issue." Such methodological problems, however, are rarely presented in articles reporting on survey

findings. Clearly, what occasioned them was the enormous differences in Gallup's results which had been reported on the front-page of the Times on November 2, while the Times' study was still in the field and their own data published on November 5. Since the Times could hardly have presented such discrepant findings without commenting on them, it followed the unusual procedure of devoting considerable space to explaining the possible sources "for an unusually high margin of error" in all opinion surveys.

As a final point, it is also important to note that the opinions of the public even those expressed a few days before an event takes place may be sharply different from their reaction to the accomplished fact particularly if it is one initiated by an important leader figure such as the President. George Gallup pointed this out at the end of the 1930's with respect to reactions to foreign and military policies, conscription, etc., before and after President Roosevelt gave his views on the subject. A more recent similar development has been discussed by Rosenberg and his colleagues:

The role of Presidential prestige and the willingness of the American public to go along with Presidential activities once he has acted can be seen rather clearly in the reaction to President Nixon's decision to send troops into Cambodia at the end of April, 1970. As Presidential actions go, this was perhaps one of the least popular actions of the Indo-China war. Yet the data are most striking.

On the eve of the Cambodian invasion, the Harris Poll asked a sample of the American population how they would feel about the commitment of American troops to Cambodia. Only 7 percent favored sending troops while 59 percent rejected such a commitment. (Twenty-three percent approved the sending of advisers and the rest were undecided.)

What happened a few days later when the President did commit troops?...Despite the very small number who favored sending American troops to Cambodia before the President did just that and despite the skepticism and apprehension of a very large majority after they were sent, when the Harris Poll asked whether Nixon was right in sending troops, more said "yes" than "no." Fifty percent agreed with Nixon's decision while 43 percent had doubts

These data vividly illustrate the prestige of the President when he acts and the malleability of American opinion. The wide gap between the 7 percent who favored sending troops before they were sent and the 50 percent who approved the President's decision after he had decided to send the troops is a measure of the support he can arouse.... (1970: 26-28).

This discussion is not intended to undermine opinion research, but rather to emphasize the need for greater care. Most issues are quite complicated and as noted earlier, require a number of questions to explore the nature of views, including contradictory opinions on the same issues, which respondents have. Many matters asked of people are basically of little concern to them; they often know little and care less, yet they answer questions. A study of attitudes toward the John Birch Society in the 1960's found that when those who said they approved of the organization were also

asked whether they thought it was a leftist or rightist group, that one-third of them described the Birch Society as leftist. But most polls which inquired about attitudes to Birch did not try to discover what image, if any, respondents had of the Society. What is one to make of the fact as found in a 1974 survey that over a third of those who said they preferred George Wallace to Richard Nixon or George McGovern agreed to the item, "I would not vote for a right-winger."

Public confidence in the polls largely rests on the fact that pre-election surveys have a very good track-record with respect to anticipating the reactions of the electorate in choosing between two men, occasionally three in Presidential contests. Fortunately, for the image of election polls, as noted earlier, no one has bothered to check up on their record in forecasting elections in primaries, mayoralty or statewide contests, one which is much less good up to literally yesterday, the 1975 mayoralty elections.

Yet all is not lost, for it seems clear that if our concern is to understand the factors associated with different views or behaviors, rather than their absolute magnitude, that relationships are generally consistent and reliable. Scales of opinion preferences, buying behavior, media habits can be related to demographic and attitudinal variables so we can know what kinds of people are social conservatives or own Volvos. It is possible to analyze what kinds of people will vote for a particular candidate, or what his image is, or that of a product, among different groups in the population and pollsters can specify how hard or soft commitment to different views are. The same questions may be repeated over time, in order to estimate the direction and approximate magnitude of changes in views or behavior.

But it must be reiterated, it is not really possible to know the opinion of the public on most issues, since there is no such opinion. There is at best usually a set of predispositions among many, but locating these do not enable us to predict behavior, or subsequent poll results with any precision.

Given all this, my counsel to all involved in survey research is humility, caution, complexity. They may reject it as ruinous to business, but I do not think making the client aware of the limits of survey findings will lose business. For people, whether businessmen, politicians, editors or the more politicized segment of the public, have an insatiable, uncontrollable need to know something, anything about public reactions. They will pay for research, for reports, for articles, even if they are reminded of the weak needs on which the conclusions rest. And by reemphasizing the instability of many attitudes and preferences of the public, we may help restore the role of judgment, of active leadership in policy, in decision-making rather than the pattern of follower-leadership which is currently so prevalent.

#### REFERENCE

Milton Rosenberg, Sidney Verba and Philip Converse, Vietnam and the Silent Majority (New York: Harper & Row, Publishers), 1970.

# ATTITUDINAL INFLUENCE ON RETAIL PATRONAGE BEHAVIOR

Masao Nakanishi, University of California, Los Angeles

## Abstract

Traditional research on retail patronage behavior which relies on the gravitational models has long ignored attitudinal factors from its design. The purpose of this study is to demonstrate that retail patronage behavior may be significantly explained by shopper attitude scale scores that are derived from paired comparison data and incorporated into a gravitational formulation.

## Introduction

In recent years much of the effort in consumer research has tended to focus on products and brands, while retail store patronage behavior has not managed to generate the research it perhaps deserves. Traditionally, research into retail patronage behavior has used models which are analogous to the model of gravitation in physics (Reilly, 1929; Huff, 1962). These approaches typically use shopping center (or store) size as the surrogate measure of center (or store) "attraction," and physical distance or travel time as the surrogate measure of "resistance" which shoppers must surmount in reaching their destination.

While these gravitational models have performed reasonably well in predicting the spatial behavior of shoppers, the validity of using physical or functional distance measures and selling space as explanatory variables has come under criticism. It has been argued (Mittelstaedt, *et al.*, 1974) that cognitive (or cognized) distance reflects more closely both the physical and evaluative dimensions which influence shopper behavior. The use of selling floor space to represent retail "attraction" has been questioned by others. Kotler (1971) argues that different images, accessibilities, and prices of retail centers create a basis for differential store preference. LaLonde (1962) tested an extension of the gravitational model which incorporates various factors and found that individual store size per se does not have the great influence claimed on drawing power. Nakanishi and Yamanaka (1975) also found that shopping center size and travel time together do not fully account for the drawing power of retail centers.

The purpose of this study is to examine the relationship between retail store patronage and consumer attitudes on such attributes as accessibility, merchandise quality and store atmosphere. The scale values derived from paired comparison data are used to test whether these variables account for store choice in a shopping trip.

## Method

### Data

The location selected for data collection was a small community in a suburban area of Osaka, Japan. A survey of the patronage patterns of two superstores (combination of discount house and supermarket) was conducted during spring of 1973. Store A with a gross selling space of 35,000 sq. feet had opened four years prior to the study, and had been enjoying a dominant position among retail stores in the community. Store

B had opened a year prior to the study within 650 yards of Store A and with a gross selling space nearly one and one-half times that of Store A.

A cluster sample of 400 housewives were selected from the residential area surrounding the two superstores. The area was divided into 14 subareas based on natural boundaries and distance from the stores. See Figure 1. From these 14 areas 80 interviewing locations were selected and five interviews conducted per cluster. Interviews were conducted in the homes and each respondent (housewife) was asked, among other questions, how often she shopped at each superstore per month. In addition each respondent was asked to compare the two stores on each of the following attributes:

1. ease of travel and access to the store
2. the quality of merchandise<sup>1</sup>
3. quantity and assortment of merchandise
4. lowest price
5. demeanor of store personnel
6. store atmosphere (lighting, cleanliness, etc.)
7. feeling of comfort and ease while shopping.

Each of these items (except 1) was repeated for three major merchandise categories: Clothing, Food, and Sundries. From the shopping frequency data, each store's "share of shopping trips" figure was computed for each merchandise category within each area. The responses for items 1 through 7 above were summarized in each subarea by calculating the proportion of respondents who rated Store A higher than Store B with regard to each attribute for each merchandise category.

### Model

Although the usual gravitational model does not accommodate the type of data obtained in this study (i.e., paired comparisons), a scaling model has been developed to integrate paired-comparison data into a gravitational-type model (Nakanishi & Cooper, 1975). There are three basic assumptions to this scaling model.

1. The shopper's attitude toward a store may be measured along several attitude dimensions (or attributes). Let  $X_{kij}$  be the underlying attitude measure of a shopper from area  $i$  toward store  $j$  along the  $k$ th attribute dimension. Assume  $-\infty < X_{kij} < \infty$ .
2. The probability that a shopper from area  $i$  rates store  $j$  over store  $j'$  in a pair-wise comparison along the  $k$ th attribute dimension is given by

$$\pi(k)_{ijj'} = \frac{e^{\gamma_k X_{kij}}}{e^{\gamma_k X_{kij}} + e^{\gamma_k X_{kij'}}} = \frac{1}{1 + e^{\gamma_k (X_{kij'} - X_{kij})}}$$

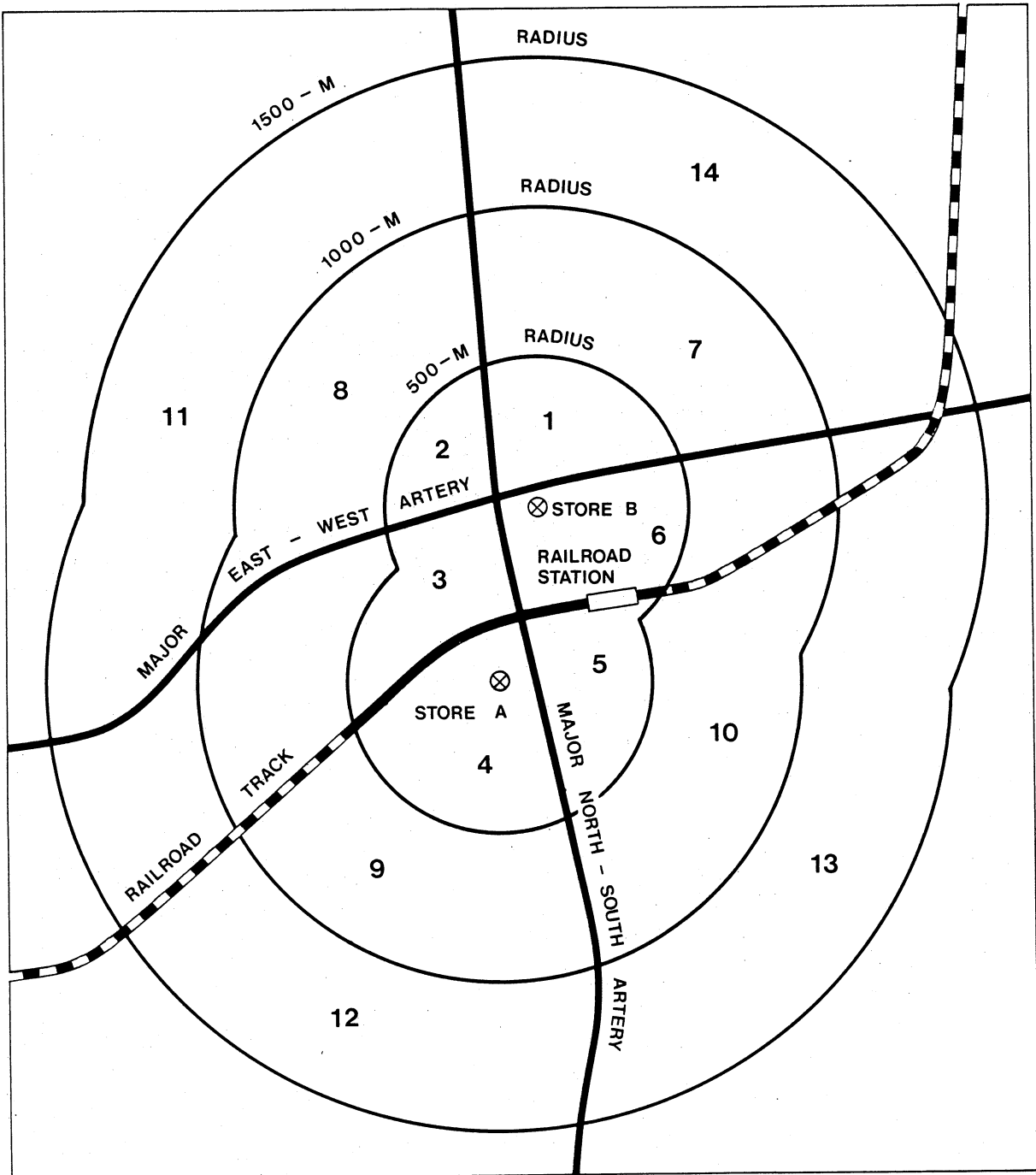
where  $\gamma_k$  is a scale parameter.

3. The probability that a shopper from area  $i$  chooses store  $j$  in a shopping trip is given by

<sup>1</sup>It may be noted that there was some difference in the wording of item 2 merchandise categories. The wording was "sense (or taste)" for Clothing, "freshness" for Food, and "quality" for Sundries.

**FIGURE 1**

**SIMPLIFIED MAP OF THE STUDY AREA**





$$\pi_{ij} = \frac{\exp\left(\sum_{k=1}^q \alpha_k X_{kij}\right)}{\sum_{j=1}^m \exp\left(\sum_{k=1}^q \alpha_k X_{kij}\right)} \quad (1)$$

where  $\exp(\cdot)$  is an exponential function,  $m$  is the number of stores,  $q$  is the number of attributes, and  $\alpha_k$  is the parameter of shopper sensitivity toward the  $k$ th attribute.

Note that the last assumption--equation (1)--is an extension of the traditional gravitational model, incorporating attitude measures as explanatory variables.

The three assumptions permit one to use the paired comparison data for estimating the parameters of the model. First, assumptions 1 and 2 lead to the estimates of the attitude scale score assigned to store  $j$  by shoppers in area  $i$  with respect to attribute  $k$ . Let  $P_{(k)ijj'}$  be the observed proportion of respondents in area  $i$  who rate store  $j$  over  $j'$  with respect to the  $k$ th attribute. ( $P_{(k)ijj'} = 1/2$  if  $j = j'$ , by assumption 2. The desired attribute scale score is given by

$$P_{(k)ij} = \frac{1}{m} \sum_{j'=1}^m \log \left( \frac{P_{(k)ijj'}}{1 - P_{(k)ijj'}} \right) \quad (m = 2). \quad (2)$$

It must be pointed out that  $P_{(k)ij}$  is an estimate of  $\gamma_k(X_{kij} - \bar{X}_{ki})$ , where  $\bar{X}_{ki}$  is the mean of  $X_{kij}$ . Next, from assumptions 2 and 3, the following regression equation is derived.

$$P_{ij} = \sum_{k=1}^q \beta_k P_{(k)ij} + \epsilon_{ij} \quad (i = 1, 2, \dots, 14; q = 7) \quad (3)$$

where:  $P_{ij} = \log(P_{ij}/\bar{P}_i)$ ,

$P_{ij}$  = observed share of shopping trips for store  $j$  in area  $i$  ( $\bar{P}_i$  is the geometric mean of  $P_{ij}$  over  $j$ ),

$\beta_k = \alpha_k/\gamma_k$ ,

$\epsilon_{ij}$  = stochastic error term.

See the Appendix for derivation of (2) and (3). This last regression model was fitted to and the  $\beta_k$ 's were estimated for each merchandise category (Clothing, Food, or Sundries) separately.

## Results

Table 1 gives the regression results based on model (3). The regression coefficients are given in the standardized form (i.e., "beta coefficient") to take account of the non-equality of variances among the attribute scale estimates.<sup>2</sup> The share of shopping trips for clothing

items is most strongly affected by the atmosphere of store, but negatively affected by store personnel and merchandise quality. The ease of travel and access has a relatively minor influence on the share of shopping trips for clothing items. For food items, the share of shopping trips is most significantly affected by the ease of travel and access and the lowness of price, but negatively affected by the atmosphere of the store. For sundries, the ease of travel and access again is the most significant attribute and other factors have virtually no effect. These findings are also supported by the regression coefficients of the "best" equation for each merchandise category, which are formed by deleting non-significant variables from the full equation.

At this point, some counter-intuitive results, such as the negative coefficients for personnel and merchandise quality for clothing items, may be questioned. One may suspect that the results are unreliable due to multicollinearity, since scale scores for different attribute dimensions may be highly correlated with each other, owing to a "halo" effect. Indeed, the  $P_{(k)ij}$ 's are highly correlated with each other, with product-moment correlation coefficients ranging as high as .95. In order to remove multicollinearity from the data, the attribute scale scores were factor-analyzed. Since there were slight differences in wording of the questions between merchandise categories, scale scores for each merchandise category were factor-analyzed separately. Because the purpose of analysis is to produce orthogonal dimensions that are parsimonious, the covariance matrix, rather than the correlation matrix, was factored. Resulting factors were Varimax-rotated. For all merchandise categories, two factors were sufficient to cumulatively explain more than 90% of over-all variance. Table 2 gives the factor analytic results.

The pattern of factor loadings are highly similar across merchandise categories. Factor 2 for clothing and food and Factor 1 for Sundries are clearly related to the ease of travel and access. Factor 1 for clothing and food represent generally the impression of the store and merchandise, while Factor 2 for sundries is associated with the impression of the store and personnel. Factor scores for these two factors are then substituted for raw scale scores in regression model (3).

The regression results using factor scores as independent variables are given in Table 3. The regression coefficients are again standardized. The factor that represents the ease of travel and access (Factor 2 for clothing and food and Factor 1 for sundries) is more important than the other factor in determining the share of shopping trips for food and sundries, but not for clothing. The impression on the store and merchandise is the most significant factor for clothing, but relatively less important for food and sundries. These results generally confirm prior expectations on shopper behavior. The explanatory power of the factor scores, measured by  $R^2$ , is high for food and sundries, but there is a significant reduction in the case of clothing (from  $R^2 = .902$  to  $.708$ ). Judging from the high percentage of

<sup>2</sup>Since estimated attribute scale score,

$P_{(k)ij} \cong \gamma_k(X_{kij} - \bar{X}_{ki})$ , regression model (3) gives the estimate of  $\beta_k = \alpha_k/\gamma_k$ , but not that of  $\alpha_k$  per se. (See the Appendix for details.) But, when standardized, the regression estimates of  $\beta_k$  becomes equivalent to that of  $\alpha_k$ . To show this, let  $\alpha_k^*$  and  $\beta_k^*$  be the standardized regression coefficient of  $\alpha_k$  and  $\beta_k$ , respectively. Also let  $\sigma_p(k)$ ,

$\sigma_{x(k)}$ , and  $\sigma_p$  be the standard deviation of  $P_{(k)ij}$ ,  $X_{kij}$ , and  $P_{ij}$ , respectively. Then,

$$\begin{aligned} \beta_k^* &= \frac{\sigma_p(k)}{\sigma_p} \cdot \beta_k = \frac{[\gamma_k \cdot \sigma_{x(k)}]}{\sigma_p} \cdot \beta_k \\ &= \frac{\sigma_{x(k)}}{\sigma_p} \cdot (\gamma_k \beta_k) = \frac{\sigma_{x(k)}}{\sigma_p} \cdot \alpha_k = \alpha_k^* \end{aligned}$$

TABLE 1  
RAW SCORE REGRESSION RESULTS

Attributes	Standardized Regression Coefficients					
	Clothing		Food		Sundries	
	Full Equation	Best Equation	Full Equation	Best Equation	Full Equation	Best Equation
EASE OF TRAVEL	.41	.60 <sup>a</sup>	1.20 <sup>a</sup>	1.12 <sup>a</sup>	.81 <sup>a</sup>	.95 <sup>a</sup>
QUALITY <sup>d</sup>	-.33	-.31 <sup>c</sup>	.15	--	-.00	--
ASSORTMENT	.12	--	-.16	--	-.05	--
LOW PRICE	-.83	--	.58 <sup>c</sup>	.46 <sup>b</sup>	.01	--
PERSONNEL	-.78 <sup>b</sup>	-.67 <sup>b</sup>	-.05	--	-.13	--
ATMOSPHERE	1.47 <sup>a</sup>	1.44 <sup>a</sup>	-.53	-.49 <sup>b</sup>	.27	--
FEEL OF AT EASE	.79	--	-.10	--	.13	--
R <sup>2</sup>	.912	.902	.963	.956	.927	.900
F-Ratio	10.3 <sup>a</sup>	23.0 <sup>a</sup>	26.2 <sup>a</sup>	79.3 <sup>a</sup>	12.7 <sup>a</sup>	117.5 <sup>a</sup>
degrees of freedom	(7,7)	(4,10)	(7,7)	(3,11)	(7,7)	(1,13)

- a. significant at the  $\alpha = .01$  level.
- b. significant at the  $\alpha = .05$  level.
- c. significant at the  $\alpha = .10$  level.
- d. different wording was used for each merchandise category.

TABLE 2  
FACTOR ANALYSIS RESULTS

Attributes	Orthogonally Rotated Factor Loadings					
	Clothing		Food		Sundries	
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2
EASE OF TRAVEL	.18	1.13	.13	1.13	1.14	.12
QUALITY	.14	-.20	.79	.27	.20	.23
ASSORTMENT	.42	-.33	.88	.17	.05	.23
LOW PRICE	.58	-.28	.65	.19	.00	.14
PERSONNEL	.31	.07	.24	.39	.12	.37
ATMOSPHERE	.55	-.09	.53	.40	.28	.40
FEEL OF AT EASE	.91	-.10	.41	.41	.13	.34
Cumulative Variance Explained	95%		95%		91%	

raw score variance explained by the two factors, the significance of the negative regression coefficients for the merchandise quality and store personnel attributes for clothing in the raw scores regression results is probably spurious.

#### Summary and Discussion

This study was conducted to see if attitudinal factors

may effectively be brought into the traditional gravitational model of shopper behavior. The predictive ability of estimated attribute scales scores (generated from paired comparisons) were generally high, explaining more than 90% of the variance in the share of shopping trips figures for different subareas. As expected, the ease of travel and access is the prime consideration in choosing a store for a shopping trip for food and sundries; this attribute is less important for choosing a store for clothing items. For other attri-

TABLE 3

## FACTOR SCORE REGRESSION RESULTS

Factors	Standardized Regression Coefficients		
	Clothing	Food	Sundries
FACTOR 1	.73 <sup>a</sup>	.16 <sup>c</sup>	.93 <sup>a</sup>
FACTOR 2	.43 <sup>b</sup>	.95 <sup>a</sup>	.23 <sup>b</sup>
R <sup>2</sup>	.708	.925	.919
F-Ratio	14.6 <sup>a</sup>	73.6 <sup>a</sup>	68.3 <sup>a</sup>
degree of freedom	(2,12)	(2,12)	(2,12)

- a. significant at the  $\alpha = .01$  level.  
b. significant at the  $\alpha = .05$  level.  
c. significant at the  $\alpha = .10$  level.

bute dimensions, the conclusion is less clear. Regression results using raw attribute scores indicate that store atmosphere is important for the choice of store for clothing items and that price is important for food items. But the negative coefficients for merchandise quality and store personnel in the case of clothing and that for store atmosphere in the case of food are intuitively unappealing. Regression results using factor scores show that the factor representing the impression of store and merchandise results in a positive coefficient for all merchandise categories and a particularly significant one for clothing.

Thus the study clearly demonstrated the usefulness of attribute scale scores from paired comparison data in predicting the drawing power of stores, but there are some unresolved issues. First, in order for the researcher to be able to determine relative importance of attributes, it is necessary to select attribute dimensions which are as orthogonal as possible to avoid the multi-collinearity problem. How one should phrase paired comparison questions to maintain the orthogonality of scale scores has not become clear. Second, to the store manager, it is not enough to know that his store suffers from a poor image; he needs to know how a poor image may be changed. It is therefore necessary to relate each attribute dimension to a set of store and location characteristics. Of course, some of the attributes are not under the control of the store manager. A more global change in the image of the entire chain may be called for. Finally, other methods for measuring shopper attitudes toward stores should be explored. For example, one may ask each respondent to rate each store on interval-scaled attitude scales. This individual-level approach, however, suffers from the fact that individual attitude scores will somehow have to be aggregated for each area since the gravitational formulation requires that the scale score for each store be determined for each area. The effect of such aggregation has not been investigated so far. The techniques used in this study, aside from its ease of administration, handles the aggregation problem in a natural manner.

## Appendix

It may be shown (Nakanishi and Cooper, 1975), that, based on the three basic assumptions described in the text and the assumption that the shoppers from each area are homogeneous in terms of the response to store attributes, we may estimate the parameters of the model, as well as the scale score for each store by attribute.

Let  $P_{(k)ijj'}$  be the observed proportion of the respondents in area  $i$  who rate store  $j$  over  $j'$  with respect to the  $k$ th attribute dimension. Since  $P_{(k)ijj'}$  is an estimate of  $\pi_{(k)ijj'}$ , from assumption 1 and 2, we have

$$P_{(k)ijj'} = \log \left( \frac{P_{(k)ijj'}}{1 - P_{(k)ijj'}} \right) \\ = \gamma_k (X_{kij} - X_{kij'}) + \eta_{kijj'}$$

where  $\eta_{kijj'}$  is the error term representing the sampling variation accruing to  $P_{(k)ijj'}$ . By averaging  $P_{(k)ijj'}$  over  $j'$ , we have

$$P_{(k)ij} = \frac{1}{m} \sum_{j'=1}^m P_{(k)ijj'} \quad (A) \\ = \frac{1}{m} \sum_{j'=1}^m [\gamma_k (X_{kij} - X_{kij'}) + \eta_{kijj'}] \\ = \gamma_k (X_{kij} - \bar{X}_{ki.}) + \bar{\eta}_{kij}$$

where  $\bar{X}_{ki.}$  and  $\bar{\eta}_{kij}$  are means of  $X_{kij}$  and  $\eta_{kijj'}$ , respectively. This shows that  $p_{(k)ij}$  is an estimate of  $\gamma_k (X_{kij} - \bar{X}_{ki.})$ , and its accuracy improves with the increase in the sample size per area and the number of alternative stores,  $m$ .

Furthermore, it may be shown that a simple transformation of equation (1)--assumption 3--yields a linear expression in parameters  $\alpha_k$ .

$$\log \left( \frac{\pi_{ij}}{\pi_{i.}} \right) = \sum_{k=1}^q \alpha_k (X_{kij} - \bar{X}_{ki.}) \quad (B)$$

where  $\tilde{\pi}_{i.}$  is geometric mean of  $\pi_{ij}$  over  $j$ . Let  $P_{ij}$  be the observed proportion of shoppers from area  $i$  choosing store  $j$  in a shopping trip. Then

$$P_{ij} = \log \left( \frac{P_{ij}}{\tilde{P}_{i.}} \right) = \sum_{k=1}^q \alpha_k (X_{kij} - \bar{X}_{ki.}) + \epsilon'_{ij} \quad (C)$$

where  $\epsilon'_{ij}$  is the stochastic error term. Since

$$\gamma_k (X_{kij} - \bar{x}_{ki}) = p_{(k)ij} - \bar{\eta}_{kij},$$

we may write (C) as

$$p_{ij} = \sum_{k=1}^q \beta_k p_{(k)ij} + \epsilon_{ij} \quad (D)$$

where  $\beta_k = \alpha_k / \gamma_k$  and  $\epsilon_{ij} = \epsilon'_{ij} - \sum_{k=1}^q \beta_k \bar{\eta}_{kij}$ . (D) may be treated as a regression equation and  $\beta_k$  may be estimated by a least squares method. Since  $p_{(k)ij}$  contains an error term due to sampling, equation (D) is an "errors in variables" formulation.<sup>3</sup> However,  $\bar{\eta}_{kij}$  approaches zero as the sample size per area approaches infinity. Thus, the ordinary least squares (OLS) estimates of  $\beta_k$  become consistent as the sample size per area approaches infinity. For further properties of the OLS estimates see Nakanishi and Cooper (1974).

#### References

Huff, David L., "Probabilistic Analysis of Consumer Spatial Behavior," W.S. Decker, ed., Emerging Concepts in Marketing. Chicago: American Marketing Association, 1962, 443-61.

Kotler, Philip, Marketing Decision Making: A Model Building Approach. New York: Holt, Rinehart and Winston, 1971, 316-20.

LaLonde, Bernard J., "Differentials in Supermarket Drawing Power," Marketing and Transportation paper 11. East Lansing: Bureau of Business and Economic Research, Michigan State University, 1962.

Mittelstaedt, Robert, William W. Curtis, Sanford L. Grossbert and Robert D. Rogers, "Psychological and Evaluative Dimensions of Cognized Distance in an Urban Shopping Environment," in Combined Proceedings. Chicago: American Marketing Association, 1974, 190-3.

Nakanishi, Masao, and Lee G. Cooper, "Parameter Estimation for a Multiplicative Interaction Model--Least Squares Approach," Journal of Marketing Research, 11 (August 1974), 303-11.

\_\_\_\_\_ and \_\_\_\_\_, "A Scaling Approach to the Synthesis of Choice Models and Multiattribute Attitude Models" Working Paper (in preparation). Los Angeles, 1975.

\_\_\_\_\_ and Hitoshi Yamanaka, "Measurement of Drawing Power of Retail Centers: Regression Analysis," Working Paper Series No. 30. Los Angeles: Center for Marketing Studies, University of California, Los Angeles, 1975.

Reilly, William J., "Methods for the Study of Retail Relationship," University of Texas Bulletin, No. 2944 (November 22, 1929).

<sup>3</sup> See J. Johnston, Econometric Methods (New York: McGraw-Hill, 1963), especially pp. 148-76.

## CUSTOMER STIMULATION NEEDS AND INNOVATIVE SHOPPING BEHAVIOR:

### THE CASE OF RECYCLED URBAN PLACES

S. L. Grossbart, University of Nebraska-Lincoln  
R. A. Mittelstaedt, University of Nebraska-Lincoln  
S. P. DeVere (student) University of Nebraska-Lincoln

#### Abstract

The recent phenomenon of recycling urban facilities into shopping complexes is linked with innovation types and the concept of individual optimal levels of stimulation. Recycled facilities are viewed as retail innovations offering a multitude of stimulation possibilities. Results indicate support of the hypotheses that high sensation seekers are more innovative and exhibit higher interstage conversion in their innovation decision processes.

Diversity is the word that perhaps best describes the physical environment of urban areas. At the city's core, diversity is maximized. Although other portions of the urban area may be physically homogeneous, the inner city's central shopping area is likely to offer a varied physical and behavioral environment. The juxtaposition of planes and surfaces, the movements of people and vehicles and the great variety of goods and services offer a set of stimulating perceptual and experiential possibilities. Here a mixture of life-style possibilities, including work, living, leisure and shopping, and a wide range of perceptual opportunities are presented to inhabitants.

As cities face the problems of maintaining the viability of their central shopping areas they often consider the alternative of tearing down and replacing existing structures. However, to preserve the urban core and maintain its central shopping areas, many cities are choosing to restore existing buildings for use as retail facilities. Architects, retailers and city planners are all urging this choice for reasons of nostalgia, conservation, and profit. Seemingly decrepit buildings which once contained warehouses, offices, governmental agencies and the like are being converted into retail centers offering varied collections of shops and services. Older districts composed of exceptional but now deteriorating buildings are being rediscovered as places capable of sustaining retail environments which can draw an increasing number of customers (Wolf, 1974). In some instances the present "recycling" activity stems from an effort to save historical sites (Harney, 1974). To some individuals, recycling provides the added advantage of preserving familiar structures and institutions. In this way historic linkages to individual structures, whole neighborhoods or even districts can be retained. In addition, the energy crisis and the diffusion of environmental awareness have decreased the perceived relative advantage of highrise structures over existing buildings and thereby have given additional impetus to the recycling trend (Harney, 1974).

The range of recycling projects is vast. In some instances single buildings have been recycled. An old warehouse in Minneapolis was converted into a complex of shops and offices; a turn of the century terminal station in Chattanooga, Tennessee has been recycled into the site of individual shops and restaurants; Salt Lake City's trolley barns have been converted into "Trolley Square," a large entertainment and shopping complex, containing ninety separate businesses ranging

from high-fashion shops to movie houses (Fortune, 1975); a parking garage in Cambridge, Massachusetts has been adapted to an attractive series of retail shops (Architectural Record, 1974). Some recycling efforts involve larger areas. Larimer Square in Denver, Old Market Square in Houston, Ghiradelli Square and the Cannery in San Francisco and Canal Square in the Georgetown area of Washington D. C. are examples of projects that have recycled complete districts into attractive areas that include shopping facilities (Mertes, 1971-72, Osman, 1974).

In investigating behavioral reactions toward recycled facilities, researchers would seem advised to treat them as retail innovations. To the extent that the stores themselves are new, this conclusion seems obvious. However, recycled facilities are also likely to be perceived as innovations insofar as they disrupt existing consumption patterns (Robertson, 1971), and present a set of consumption related opportunities in a context which is not necessarily compatible with the range of typical consumer values or experiences (Rogers and Shoemaker, 1971).

The present study focuses on customer reactions toward a block long section of empty warehouses in a Midwestern city. In the discussion which follows an innovation decision framework is developed to identify the psychological processes leading to acceptance of new recycled facilities. In addition, an argument is set forth and tested regarding the customer types likely to be differentially attracted by the stimulus potential of innovative recycled facilities.

#### Decision Behavior

The behavior leading to innovative or adoptive decisions is typically envisioned as a multiple stage process (Rogers and Shoemaker, 1971; Robertson, 1971). Rogers and Shoemaker's (1971) four function knowledge-persuasion-decision-confirmation model rests on a conceptualization which distinguishes between reactions to ideational and object type innovations. Although Rogers and Shoemaker consider innovations with a physical presence to have both an idea and object component, they specify the existence of a symbolic decision only for purely ideational innovations, such as an ideology. In contrast, for physical innovations with an object component, e.g., products, the ultimate decision is seen as an observable act.

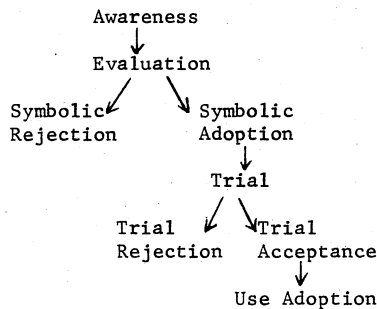
In examining customer decision behavior regarding retail facilities, the application of a nonsymbolic decision framework raises two critical problems. First, stores clearly have a symbolic character (Kotler, 1973-74) and it is likely that the acceptance of this ideational component of facilities would be a requisite for their confirmed adoption into a customer's set of acceptable retail alternatives. Second, since the use of many types of facilities is infrequent or sporadic, the absence of a symbolic as well as an overt behavioral measure of adoption would lack valuable information useful for tracking diffusion processes. In many instances the criterion for adoption would remain

operationally nebulous.

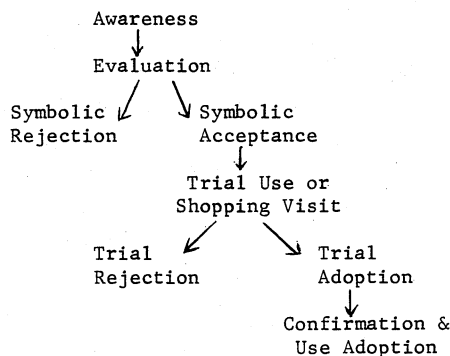
A scheme which incorporates these ideas can be based on the work of Klonglan and Coward (1970). Their decision process model incorporates both symbolic and action forms of adoption (Figure 1a). In this model awareness precedes evaluation, which may produce either symbolic acceptance or rejection of the idea of the innovation. While symbolic rejection is not irrevocable, it does mean, barring a reversal, that customers have concluded that the store or stores are not suitable to meet their needs. Symbolic acceptance is a requisite for trial but does not assure engagement in trial. Customers may accept the idea of the facilities but, for some reason, delay trial (Figure 1b). For example, they may conclude that they are not presently "dressed appropriately," that "it is too late today," that they do not presently "need any \_\_\_\_\_," etc. The common factor here is the presence of a symbolic act of acceptance and the postponement of an overt act. Following trial, the customer either rejects further interaction with the facility or confirms the shopping decision. This confirmation, while reversible, represents the adoption of the retail facility. Trial rejection, while also revocable, provides a basis for revoking the initial decision to shop or use a facility. The opportunity for this type of rejection allows a customer to reject stores on the basis of little information and direct experience, which is likely to be the condition for new facilities, or to reject stores on the basis of considerable information and direct experience, which is the typical condition once a diffusion effect (Rogers and Shoemaker, 1971) takes over. The operative conditions depend on the timing of the customers awareness-knowledge as well as the customers need for information prior to trial.

FIGURE 1

(a)<sup>1</sup> A Two Phase Model of the Adoption Process



(b) A Modification of The Two Phase Model of The Adoption Process For New Retail Facilities



<sup>1</sup>Adapted from Klonglan and Coward "The Concept of Symbolic Adoption: A Suggested Interpretation," *Rural Sociology*, 35, (March 1970), p. 80.

The modified framework presented in Figure 1b provides the user with two types of operational opportunities. The propensities of alternative customer group to gain knowledge, evaluate, symbolically accept, etc. new facilities can be compared to determine alternative market segments. In addition, the tendency for different customer types to move from one decision state to another, e.g., from trial to confirmation, can be examined. The likelihood of moving from one decision state to another is referred to as a conversion probability. (For a similar conceptualization see Engel, Kollat and Blackwell, 1973, p. 520-21.) Conversion analysis can give a macro-picture of the manner in which decisions are made by alternative customer segments and simultaneously suggest how these groups might best be dealt with in tactical terms.

Before applying this decision behavior framework to a situation involving recycled retail facilities it is necessary to discuss innovative behavior as a manifestation of stimulation needs, the operationalization of such stimulation needs, and the rationalization for linking the need for stimulation to the acceptance of recycled facilities.

#### Customer Stimulation Needs and Decision Behavior

Customers use of recycled facilities provides them an opportunity for novel experiences and customer acceptance of these new facilities represents an innovative act. It has been suggested elsewhere that reactions to novel consumption experiences are related to individual optimal levels of stimulation. Individuals are said to seek to maintain an optimal level of stimulation, with departures from optimality leading to behavior intended to reestablish an optimal intermediate range of stimulation (Jones, 1969; Kish and Busse, 1968; Zuckerman, 1967). Although similar conclusions have been derived from different theoretical psychological frameworks empirical evidence clearly supports the existence of a motivation for stimulation (Maddi, Propst and Feldinger, 1965; Hunt, 1963; Driver and Streufert, 1965). As noted by Howard and Sheth (1969) and Venkatesan (1973) it is the novel, surprising, incongruous and ambiguous features of products which carry stimulation potential. The same can be said for retail facilities. Attributes of recycled facilities should possess the stimulus potential necessary to provide consumers with information and stimulation.

In the present context, we would expect the attractiveness of recycled facilities to vary with customers' needs for stimulation. Evidence (Zuckerman, 1971) indicates that individuals do indeed differ in their requirements for stimulation. The concept of a need for stimulation has been operationalized by Zuckerman (1964) in the form of a Sensation Seeking Scale (SSS). Application of the SSS in a variety of research settings has resulted in composite profiles of high (HSS) and low (LSS) sensation seekers (See Gorman, 1970; Kish and Donnenwerth, 1969; Kish and Busse, 1968; Norman and Fenson, 1970; Zuckerman and Schultz, 1967; Zuckerman, Nery and Brustman, 1970; Zuckerman and Link, 1968.) In comparison to the LSS, a HSS has been found to be younger, more educated, more intelligent and more risk oriented. The HSS demonstrates a greater preference for sensory variability, an attraction to new experiences, a tendency for impulsive action, and a desire for complex, changing and less structured activities and environments than a LSS. Finally, HSSs' have been found to be high in social independence and individual autonomy.

Since a HSS is likely to be interested in a variety of new experiences and since recycled facilities appear to possess many of the attributes which a HSS would

perceive as a source of stimulation, it seems reasonable to expect that HSS's would be more receptive than LSS's to such facilities. Further, because of their tendency to readily explore stimulating situations, it would be expected that the decision behavior of HSS customers, measured in terms of conversion, or movement from one decision stage to another, would differ from that of LSS customers even if each were aware of the existence of new facilities, considered shopping there, accept them symbolically, or actually visited the facilities. In short, HSS's should represent a disproportionate number of customers for such facilities and should differ from LSS's in their decision behavior. The following research hypotheses are developed in terms of the modified decision framework presented above (Figure 1b) and reflect each of these propositions.

#### Hypotheses

The proportion of HSS customers who:

- 1.1 are aware of the retail innovation
- 1.2 consider shopping at the retail innovation
- 1.3 symbolically accept the idea of the retail innovation
- 1.4 actually shop at the retail innovation
- 1.5 confirm the decision to continue shopping at the retail innovation

will exceed the proportion of LSS customers exhibiting the same behavior.

Of those customers who are aware of the retail innovation, the HSS customer will be more likely to:

- 2.1 consider shopping the retail innovation
- 2.2 symbolically accept the idea of innovation
- 2.3 actually shop at the retail innovation
- 2.4 make a confirmed decision to continue shopping at the retail innovation

than the LSS customer.

Of those customers who consider shopping the retail innovation, the HSS customer will be more likely to:

- 3.1 symbolically accept the idea of retail innovation
- 3.2 actually shop at the retail innovation
- 3.3 make a confirmed decision to continue shopping at the retail innovation

Of those customers who symbolically accept the idea of the retail innovation, the HSS customers will be more likely to:

- 4.1 actually shop at the retail innovation
- 4.2 make a confirmed decision to continue shopping at the retail innovation

than the LSS customers.

Of those customers who actually shop at the retail innovation, the HSS customers will be more likely to:

- 5.1 make a confirmed decision to continue shopping at the retail innovation

than the LSS customers.

#### Method

##### Facilities

Customer decision behavior was examined with respect to a set of retail facilities established as part of a recycling of formerly commercial warehouses. The multi-level structures were partially vacant and although the buildings were located in proximity to the downtown area, entrances did not border any existing shopping facilities. Traffic to and from the downtown does not normally flow past the buildings and in the normal course of travel activities residents are unlikely to have visual contact with the structures. The buildings themselves were renovated in two stages. The study was conducted approximately 6 months after the first stage opened and before the second group of stores was complete. The timing of the study permitted the wide potential exposure of the new facilities, the transmission of limited advertising, and the opportunity for interpersonal communication between residents (Kelly, 1967). The facilities themselves are not unlike those in other recycled structures and consist of a series of retail clothing, gift, sports, and music shops, restaurants, bars, and movie theatres. Although establishments cultivated separate identities and are operated independently, a common external facade, name, and limited joint advertising program has been developed. The recycling renovation program was the first in the city's recent history.

##### Respondents

A sample of 114 adult females in a midwestern city provided the present data base. Females were chosen because of their more probable engagement in shopping behaviors and potential exposure to retail innovations. Respondents were selected by random multi-stage sampling. The characteristics of the sample were comparable in socioeconomic and demographic terms to those of the population.

##### Sensation Seeking and Decision Behavior Measures

Form II of the general Sensation Seeking Scale GSS (Zuckerman, 1964) was administered within each respondent's residence. The scales required ten to fifteen minutes for completion and consist of twenty-two forced-choice item pairs describing alternative activities possessing different levels of stimulation.<sup>2</sup> The SSS is scored by summing the number of more stimulating activity choices.

Questions based on the modified Klomglan-Coward model (Figure 1b) were used to determine: (1) awareness of the facilities, (2) consideration of shopping, (3) symbolic acceptance or rejection, i.e., a consideration and decision whether to shop, (4) actual shopping, and (5) confirmation of the decision to continue using the facilities. This procedure provided a basis for classifying respondents as to their position in the decision process and aided in identifying apparent "holding patterns" within the process.

<sup>2</sup> For example, one pair of GSS items state:

- A. The most important goal of life is to live it to the fullest and experience as much of it as you can.
- B. The most important goal of life is to find peace and happiness.

Procedure

Respondents were classified as being a high sensation seeker (HSS) or low sensation seeker (LSS) on the basis of their summed SSS score with a division of the SSS at its median. Fifteen decision behavior measures were computed separately for each SS group. Measures of the dependent variables corresponding to the first five hypotheses were calculated by computing the proportion of each SS category indicating each of the respective decision behaviors. For the remaining ten hypotheses, measures were also developed for each SS category. In these cases, the measures indicated the proportion of customers moving from one decision state to all subsequent decision states. Thus the proportion of customers converted from awareness to consideration, symbolic acceptance, shopping, and confirmation (hypotheses 2.1 through 2.4), from consideration to symbolic acceptance, shopping, and confirmation (hypotheses 3.1 through 3.3), from symbolic acceptance to shopping and confirmation (hypotheses 4.1 and 4.2), and from shopping to a confirmed decision to continue each SS category. Tests of differences between proportions were conducted to test for significance in the case of each respective hypothesis.

Results and Discussion

The procedure outlined above resulted in the classification of 83 respondents as LSS's and 31 respondents as HSS's. A comparison of the HSS and LSS groups indicated no significant differences in education, family size, occupation of the head of the household, and family income. The HSS respondents were somewhat younger than the LSS group, but even in those age categories in which the HSS tended to cluster, they represented a minority of the total cases. Thus while a HSS is likely to be a younger individual than a LSS, there is less than an equiprobable chance that the younger individual will be a HSS.

Results are presented in Tables 1 and 2. Reference to Table 1 indicates that all differences are in the predicted direction. The HSS is more likely than the LSS to be aware of the new facilities, to consider using the facilities, to symbolically accept (or fail to symbolically reject) the facilities, to actually visit the facilities and to confirm the adoption decision

regarding the facilities. In all cases the differences are significant, thereby lending support to hypotheses 1.1 through 1.5.

TABLE 1  
Proportion of LSS and HSS Customers Reaching Each Decision Stage

Decision State	LSS (n=83)	HSS (n=31)	Z value
Awareness	.711	.968	2.95 <sup>b</sup>
Evaluation	.518	.871	3.43 <sup>a</sup>
Symbolic Acceptance	.434	.807	3.55 <sup>a</sup>
Trial Use	.241	.677	4.32 <sup>a</sup>
Confirmation	.181	.613	4.50 <sup>a</sup>

<sup>a</sup>Significant at .001 level.

<sup>b</sup>Significant at .01 level.

The conversion results appear in Table 2. This information differs from the gross percentage figures in Table 1, which reveal the likelihood that the decision state is made without indicating the nature of the decision behavior itself. The information in Table 1, for example, indicates the likelihood that all customers in a SS category will shop in the recycled facilities, while the information in Table 2 indicates the likelihood of customers shopping in the facilities if they, alternatively, are aware, consider, or symbolically accept the new facilities.

TABLE 2

Proportion of LSS and HSS Customers Converted From Each Decision Stage to Subsequent Decision Stages

Conversion To	Conversion From											
	Awareness			Evaluation			Symbolic Acceptance			Trial		
	LSS (n=59)	HSS (n=30)	Z value	LSS (n=43)	HSS (n=27)	Z value	LSS (n=36)	HSS (n=25)	Z value	LSS (n=20)	HSS (n=21)	Z value
Evaluation	.728	.900	1.87 <sup>c</sup>									
Symbolic Acceptance	.610	.833	2.14 <sup>c</sup>	.837	.926	1.07						
Trial	.339	.700	3.22 <sup>a</sup>	.465	.778	2.59 <sup>b</sup>	.556	.840	2.33 <sup>b</sup>			
Confirmation	.254	.663	3.48 <sup>a</sup>	.349	.704	2.89 <sup>b</sup>	.417	.760	2.66 <sup>b</sup>	.750	.905	1.31 <sup>d</sup>

<sup>a</sup>Significant at .001 level.

<sup>b</sup>Significant at .01 level.

<sup>c</sup>Significant at .05 level.

<sup>d</sup>Significant at .10 level.



The awareness-conversion results for both SS categories provide an initial basis for thinking about the nature of decision making by customers in each SS category. As might be expected, the probability of conversion declines with each succeeding decision state. However, it should be noted that this decline is markedly greater in the case of the LSS group. As a partial consequence, the stated differences in conversion are all in the direction predicted by hypotheses 2.1 through 2.4, i.e., the HSS who is aware of these facilities is more likely to consider, symbolically adopt, trial shop or use, and confirm such use than is the LSS who is also aware of their existence. The differences are significant in each case.

The consideration-conversion data bear these same characteristics and also possess significant differences, thereby supporting hypotheses 3.2 through 3.3. While the differences in conversion from consideration to symbolic adoption are in the direction suggested by hypothesis 3.1, they are not significant. Thus while there is no significant difference in the likelihood of symbolic adoption, given that a customer considers a set of facilities, but having considered such facilities, HSS customers are more likely than LSS customers to both try (by visit or use) and confirm the decision to adopt such facilities into their set of acceptable retail alternatives.

Once they symbolically adopt the idea of the facilities, the HSS customers are more apt, as suggested by hypotheses 4.1 and 4.2, to subject them to trial and confirm their continued use. Both differences are significant. It is interesting to note that for the LSS group the chances of a conversion to trial exceed .5 only after symbolic adoption. The probability of HSS conversion to trial exceeds .5 given only awareness and increases as these customers move through the decision process. Further, from the standpoint of awareness, consideration, or symbolic adoption, there is a less than even chance of LSS conversion to confirmation and, in contrast, a greater than even tendency for HSS conversion to confirmation.

It is only when they have actually visited the facilities that the LSS shows a confirmation tendency which exceeds equiprobability. Even in these terms, the confirmation tendency which exceeds equiprobability. Even in these terms, the confirmation probability of HSS customers is significantly greater, supporting hypothesis 5.1.

#### Implications

The recycling of urban places into shopping complexes has important economic and social implications, not only for the investors involved, but also for those concerned with the vitality of our central cities. Certainly, recycling alone will not arrest the set of forces which have endangered urban centers. At the same time successful restorations can contribute to the ultimate health of these areas. In this context, it is desirable for consumer researchers to join with both planners and retail investors in considering the likely ramifications of alternative recycling projects. The present study indicates the importance of considering questions dealing with both the types of customers attracted to such facilities and the manner in which other residents might come to consider and use the same facilities.

The findings suggests that the adoption of recycled facilities is dependent on the stimulation needs of potential customers. However, before discussing the implications of these findings, it should be noted that

the reasoning set forth does not imply that optimal stimulation levels magnify customer receptivity to any type of new retail facility. In fact, a case might be made that the opening of a new dairy store, dry cleaner, hardware store, etc. would hardly offer a source of meaningful stimulus input and would therefore be unlikely to elicit the same type of response as observed here. In this sense the opening of a new outlet of a traditional store type may represent a continuous retail innovation (Robertson, 1971) and therefore not attract the HSS. At present this conclusion is speculative. It does, however, appear necessary to limit the range of retail innovations to which inferences relating to sensation seeking might be made. It would also seem desirable to develop instruments which directly measure the stimulus impact of retail facilities.

Although traditional demographic characteristics appear to be poor predictors of which group of customers are attracted to recycled facilities, some intriguing segmentation possibilities are suggested by the present study. Support of the hypotheses indicates that those individuals who possess a higher optimal level of stimulation (HSS's) are the customers most likely to be interested in utilizing recycled facilities. That the shops, restaurants and entertainment possibilities usually located in the recycled facilities offer liberal opportunities for varied experiences and stimulation is perhaps the best intuitive explanation of the phenomenon we have observed. Individuals who do not require relatively large amounts of stimulation may in fact find satisfaction in a recycled complex, but are likely to become positive in their outlook on such facilities only after an actual trial visit. We are dealing with a group of individuals, the HSS's, who being inherently interested in recycled facilities as a source of stimulation, are likely to be sensitive or responsive to initial mass communication efforts. Alternatively, the LSS, must be provided with a stronger impetus to use of such facilities. For the LSS group, taken as a whole, it may well be that behavior (contact with the facilities) precedes the requisite degree of attitudinal change necessary for adoption. If this is the case, ultimate LSS adoption is more dependent on direct action sales promotion efforts which stimulate shopping visits rather than advertising which describes the unusual or exciting features of a recycled facility. These lines of inquiry certainly merit further study.

#### References

- "Buildings Can Be Recycled Too," Fortune, (May 1975), 192-200.
- "Conservation In The Context Of Change," Architectural Record, (Dec. 1974), 85-98.
- M. J. Driver, and S. Streufert, "The 'General Incongruity Adaptation Level' Hypothesis: An Analysis and Integration of Cognitive Approaches To Motivation," Paper No. 114 Institute for Research in the Behavioral Economic & Management Sciences, Herman C. Krannert Grad School of Industrial Administration, Purdue University Lafayette, Ind. 1965.
- J. F. Engel, D. T. Kollat, and R. D. Blackwell, Consumer Behavior, (2nd ed.) (New York: Holt, Rinehart and Winston, 1968).
- Bernard S. Gorman, "16 PF Correlates of Sensation-Seeking," Psychological Reports, 26 (1970), 741-742.
- A. L. Harney, "Adaptive Use: Saving Energy As Well As Historical Buildings," AIA Journal (Aug. 1974), 49-54.

- J. A. Howard, and J. N. Sheth, The Theory of Buyer Behavior, (New York: John Wiley & Sons, 1969), 164-165.
- J. M. Hunt, "Motivation Inherent in Information Processing and Action," in Motivation and Social Interaction-Cognitive Determinants, ed. by O. J. Harvey, (New York: The Ronald Press, 1963), 35-94.
- Austin Jones, "Stimulus Seeking Behavior," Sensory Deprivation, edited by John P. Zulzek, (New York: Appleton, Century and Crofts, 1969).
- R. F. Kelly, "The Role of Information In The Patronage Decision: A Diffusion Phenomenon," in Marketing For Tomorrow . . . Today, edited by M. S. Moyer and R. E. Vosburgh (Chicago: American Marketing Association, 1967).
- George B. Kish, and William Busse, "Correlates of Stimulus-Seeking: Age Education, Intelligence, and Aptitudes," Journal of Consulting and Clinical Psychology, 32 (1968), 633-637.
- G. B. Kish, and G. V. Donnenwerth, "Interests and Stimulus-Seeking," Journal of Counseling Psychology, 16 (1969), 551-556.
- G. E. Klonglan, E. W. Coward, "The Concept of Symbolic Adoption: A Suggested Interpretation," Rural Sociology, 35 (March 1970), 77-83.
- Phillip Kotler, "Atmospherics As A Marketing Tool," Journal of Retailing, 49 (1973-1974), 48-64.
- S. R. Maddi, B. S. Propst, and I. Feldinger, "Three Expressions of the Need for Variety," Journal of Personality, 33 (1965), 82-98.
- J. E. Mertes, "Retailers Use of Historical Sites," Journal of Retailing, 47 (1971-72), 54-64.
- Ralph D. Norman, and Judith N. Fenson, "Further Aspects of Construct Validity of the Zuckerman Sensation-Seeking Scale," The Journal of Psychology, 74 (1970), 131-140.
- Mary Osman, "Arthur Cotten Moore/Associates Of Washington," AIA Journal, (May 1974), 58-61.
- T. S. Robertson, Innovative Behavior and Communication, (New York: Holt, Rinehart and Winston, 1971).
- E. M. Rogers, and F. F. Shoemaker, Communication of Innovations, (New York: The Free Press, 1971).
- M. Venkatesan, "Cognitive Consistency and Novelty Seeking" in Consumer Behavior: Theoretical Sources, edited by S. Ward & T. Robertson (Englewood Cliffs, New Jersey: Prentice Hall, 1973), 354-384.
- P. Wolf, The Future Of The City, (New York: Watson-Guptill Publications, 1974).
- Marvin Zuckerman, "Development of a Sensation-Seeking Scale," Journal of Consulting Psychology, 28 (1964), 477-482.
- Marvin Zuckerman, and Duane P. Schultz, "Sensation Seeking and Volunteering for Sensory Deprivation and Hypnosis Experiments," Journal of Consulting Psychology, 31 (1967) 358-363.
- Marvin Zuckerman, Richard S. Neary, and Barbara A. Brustman, "Sensation-Seeking Scale Correlates in Experience (Smoking, Drugs, Alcohol, 'Hallucinations,' and Sex) And Preference For Complexity (Designs)," Proceedings, 78th Annual Convention, APA, 1970.
- Marvin Zuckerman, "Dimensions of Sensation Seeking," Journal of Consulting and Clinical Psychology, 36 (1971), 45-52.
- Marvin Zuckerman, and Kathryn Link, "Construct Validity For The S.S.S." Journal of Consulting and Clinical Psychology, 32 (1968), 420-426.

## BUYER BEHAVIOR UNDER CONDITIONS OF CROWDING: AN INITIAL FRAMEWORK

Gilbert D. Harrell, Michigan State University  
Michael D. Hutt, University of Vermont

### ABSTRACT

Environmental conditions, specifically physical density and crowding, may affect several key dimensions of retail shopping behavior. Exploratory research indicates that these forces are a salient force in the retail setting.

An emerging interdisciplinary field of inquiry, environmental psychology, has evolved which focuses on the relationship between the physical environment and human behavior. While in an early stage of development, the presence and importance of influences emanating from the environment of human behavior have been established in a number of diverse empirical studies. Clearly, environmental influences are worthy of more thorough investigation in the study of buyer behavior. The central purpose of this paper is to outline a paradigm of buyer behavior and one environmental condition, crowding.

Kotler makes an important contribution in introducing Atmospherics as a Marketing Tool (Kotler, 1974). He contends that in some cases "the place, more specifically the atmosphere of the place, is more influential than the product itself in the purchase decision" (Kotler, p. 48). Atmospherics is defined as "the conscious designing of space to create certain effects in buyers" (Kotler, p. 50). This conceptualization can be extended to include not only the created physical structure but also the atmosphere created by individuals shopping in the store. Thus, the atmosphere of the store can take on a dynamic quality, and conditions of crowding can alter the psychological atmosphere of a store to evoke different patterns of shopping behavior.

While the effects of crowding have been empirically examined by a number of researchers, exploration of the concept in the marketing setting is absent from the literature. Important trends in marketing point up the need for inquiry into the area. First, scrambled merchandising, regional shopping centers and, more recently, the super store, all require heavy concentrations of shoppers. Second, because of the growing number of working wives, available shopping hours have been cut, thus placing a heavier burden on peak shopping times, e.g., Saturdays.

When is a store "crowded"? A manager and a consumer may respond differently to this query. Stokols (1972) identifies two components of crowding: (1) a physical condition, density, involves the restriction of movement imposed by limited space, while the experiential state, crowding, encompasses the individual's perception of the restrictive aspects of limited space. The challenge for the manager is to increase density without triggering the experiential state of crowding among shoppers.

The purpose of this paper is to explore the potential relationship between crowding and buyer behavior. Our approach is to first address the theory on crowding from several disciplines and research tracks. Secondly, a paradigm of buyer behavior under conditions of crowding is proposed. Three goals are sought: (1) to stimulate thinking about crowding and other environmental

influences, (2) to provide a preliminary sketch of the possible influences of crowding on shopping behavior, and (3) to generate hypotheses for future research. The paradigm is grounded in the literature and was reinforced by the use of group depth interviews of shoppers.

### Fundamental Concepts of Crowding

Psychological or perceived crowding is a function of several environmental and individual variables in addition to the apparent lack of space. The more salient research findings suggest that crowding is a function of the situation including the difficulty of the task (Stokols, 1972); the amount of interpersonal stimulation and interaction (Desor, 1970) (Zlutnick and Altman, 1972); the individual's personal characteristics such as aggressiveness (Stokols, et al., 1973); and the individual's previous experience and expectations in the situation (Zlutnick and Altman, 1972). Moreover, the presence of these conditions results in information overload, confusion, and frustrated goal seeking (Proshansky, et al., 1972). In turn, coping or adaptation strategies are evoked and behavior is altered.

### Situational Determinants of Crowding

The obvious situational determinant of crowding is the presence of a high density of people per unit of space. Animal studies confirmed the deleterious effects of high density on behavioral patterns and population growth. For example, Calhoun (1962) reported abnormal social patterns and drastically reduced birth rates with rats confined in high density experiments. Christian, Flyger, and Davis (1960) studied deer on a naturally confined island and reported similar findings. Generally, human studies have shown that density alone may not produce aberrant human behavior in the experimental setting. Freedman, et al. (1971), found few experimental effects on task behavior by variations in experimental room size.

Since man is a cognizing and goal-directed organism dependent on interactions and exchanges in his environment, crowding occurs only when space restrictions confine goal directed behavior (Proshansky, et al., 1972). For example, high density may actually facilitate goal achievement at a social function. The "richness of environmental resources" (Zlutnick and Altman, 1972, p. 51) is definitely a factor. However, in a supermarket, high density may impede mobility and decrease shopping efficiency. The degree to which time effectiveness is altered becomes important to the shopper.

### Interpersonal Stimulation and Interaction

Desor (1972, p. 79) has stressed the interpersonal determinants of crowding by defining "being crowded" as "receiving excessive stimulation from social sources." Obviously, the amount of total stimulation from both the physical and the social environment affect the individual's perception of a situation. For example, the

Zlutnick and Altman (1972, p. 52) add another important social determinant of crowding. They point out that a feeling of crowding relates to "people's ability and inability to control their interaction with others." Crowding becomes most pronounced when interference occurs. Restricted or redirected movement due to the presence of other individuals would enhance a crowded feeling. At the same time, high density situations which allow maximum freedom of movement produce less psychological crowding. Conditions of control and movement have not been empirically studied in enough detail to specify the exact relationship to crowding.

#### Individual Characteristics

Research indicates that individual characteristics have an effect on the degree to which a particular environment is perceived as being crowded. Stokols, et al. (1973) found that aggressiveness and anxiety influenced crowding when subjects were engaged in a competitive game. Few additional probes have been made into the significance of personality traits to psychological crowding. Clearly, further research is needed.

#### Individual's Previous Experience

The psychological makeup of any individual is largely determined by previous learning. Crowding is a relative concept--environments are more or less crowded and the anchors for judgment can be somewhat elusive. The Japanese, who are accustomed to extremely high population density, perceive crowding differently than the rural American. In buying behavior it is important to determine the circumstances under which crowding is experienced by the shopper. The environment most certainly has a normative property, although no particular measure for it exists. When the norm is violated, adaptation or coping strategies become operative. By developing an understanding of these strategies the marketer could become more responsive to the needs of the consumer. The obvious alternative to crowding for the consumer is to leave the crowded store and never return. However, there are tradeoffs. To illustrate, convenience of location, price considerations, and merchandising variety may more than compensate for the crowded condition. In fact, the manager attempts to obtain high density by offering a large number of buyers a package of rewards which exceed the costs of the shopping task.

#### Adaptation and Coping When Crowding Occurs

Stanley Milgram (1970) used the concept of overload from systems theory to explain the problem of excessive social stimulation in city environments. When overload occurs, adaptations follow. While his conclusions refer to the total city environment, they offer significant insights pertinent to buyer behavior by enumerating in a generic sense several categories of responses to overload. First, less time is allocated to each stimulus input; for example, perhaps less time is allocated to comparative shopping. Second, low priority items are disregarded; thus, low priority goods may not be purchased or specials might be selectively screened out. Third, boundaries in the social system are altered so that the burden can be shifted to others in the exchange process. To illustrate, the use of credit cards alleviates the need to carry cash or write checks. Fourth, reception is blocked off by not entering the system. How many buyers avoid crowded shopping malls? Fifth, social inputs may be limited to relatively superficial forms of involvement with others. An analog to retail shopping might be the consumer's attempt to minimize contact with clerks and

other shoppers. Sixth, specialized institutions are developed to absorb inputs that might otherwise swamp the individual; for example, the yellow pages alleviate the need for some shopping trips.

#### Research Approaches to Study Crowding

Human research on crowding has employed several methodologies. The most important include: (1) aggregate level correlational studies--for example, the relationship of crime and city density (Schmid, 1960); mental illness and city density (Faris and Dunham, 1965) and city density and selected social interactions (Smith, et al., 1954); (2) experimental laboratory studies--for example, on room size and configuration and crowding (Valins and Baum, 1973), and density and crowding in task environments (Stokols, et al., 1973) (Freedman, et al., 1971); and (3) speculation and informal observation--for example, a number of informal untested hypotheses can be found (Esser, 1972), (Dubos, 1968).

Conspicuously absent from this list of methodologies is the study of individuals either in a depth interview setting or in a field environment. Insights derived both from the literature and group depth interviews of housewives were used to develop the framework of buying behavior under conditions of crowding proposed in this paper. Prior to discussing crowding in the buying context, the exploratory research methodology is described.

#### Group Depth Interviews to Assess Buying Behavior Under Conditions of Crowding

Because of the speculative nature of the research in the area and the paucity of literature on crowding and buying behavior, initial contacts with consumers were of an exploratory nature. Four group depth interviews were conducted with 46 homemakers in total. Each 90-minute interview involved 11 or 12 women who are the major food shoppers in their respective families. A trained moderator presided. The moderator functioned in a non-directive manner to facilitate group interaction and exploration. Each interview was tape recorded, transcribed, and content analyzed. The nature of the group depth approach does not allow for a highly precise quantification of results. However, four important objectives were achieved. First, the interviews allowed an in-depth probe of attitudes toward the retail shopping experience. Second, crowding emerged as a relevant dimension of buyer behavior worthy of further consideration. Third, the salient individual and social dimensions of crowding were explored. Fourth, several adaptation strategies used by shoppers in crowded store environments were isolated. Most of the discussion was anchored to the topic of food shopping, although discussion of other types of shopping often arose during the interviews. The group depth interview as a research technique is reviewed by Goldman (1962).

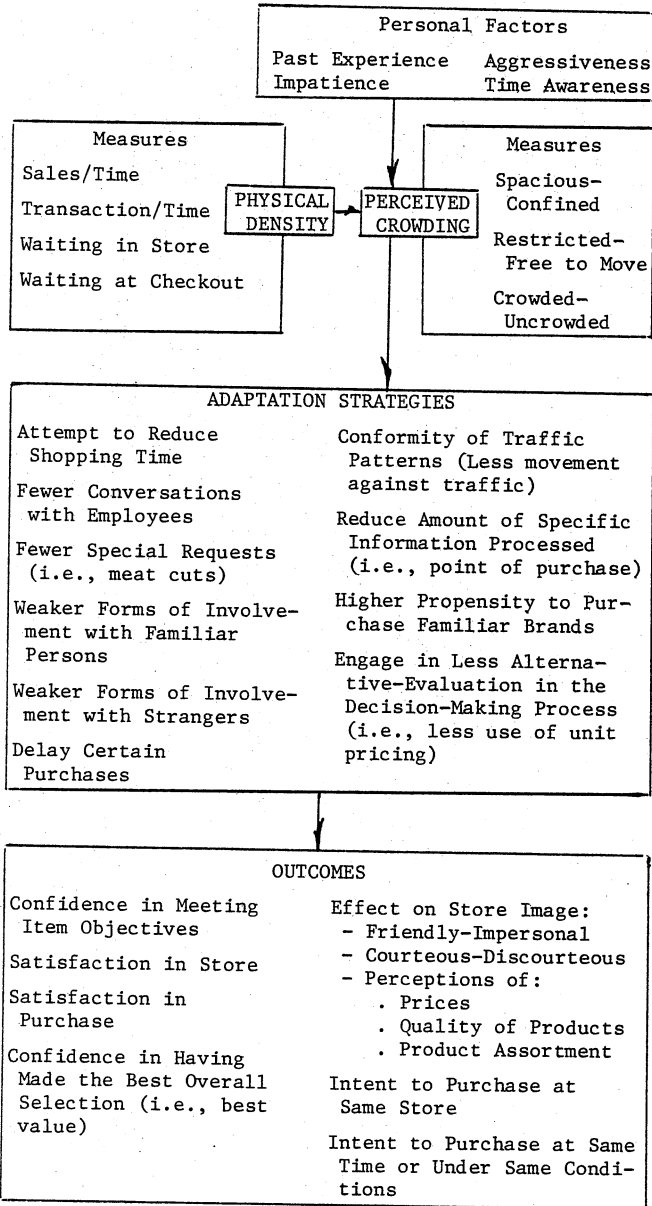
#### A Model of Supermarket Shopping Behavior Under Conditions of Crowding

Valuable insights into buyer behavior may be secured by examining crowding in a retail setting. As indicated, considerable research in crowding has been conducted in a number of disciplines and important groundwork has been established. Clearly, there is a need for an integration of relevant knowledge into a structure that may aid our understanding of retail shopping behavior. Until empirically verified, the propositions flowing from the model should be termed tentative.

A model of supermarket shopping behavior under conditions of crowding is presented in Figure 1. The three

FIGURE 1

Buyer Behavior Under Conditions of Crowding



shoppers and the physical structure and resources of the store. Possible measures of physical density include sales/time, transactions/time, and involuntary time delays both within the store shopping area and at the checkout

Psychological crowding can be operationally defined as the shopper's perceptions of the restrictive aspects of limited space. Possible measures include the dimensions of spacious-confined, restricted-free to move, and crowded-uncrowded.

Personal Factors

Psychological crowding encompasses more than high density. Several personal factors may moderate the degree to which crowding is perceived. The most important factors appear to be (1) past experience, (2) time awareness, and (3) personality characteristics such as impatience and aggressiveness. These elements were repeatedly brought out in the depth interviews. A consumer that lacks prior experience in shopping in a crowded environment or alternatively is under time pressure may be more sensitive to crowding. Likewise, character traits such as impatience and aggressiveness can render a shopper more susceptible to perceived crowding. Several buyers attributed their feelings of crowdedness to one or several personality traits that magnified the problem for them. Thus, physical density along with selected character traits and situational variables combine to determine the shopper's perception of the environment. "I feel my personal space is violated." "Confusing," "frustrating," "confining," "restricted." These comments surfaced during the group depth interviews.

Adaptation Strategies

How does a shopper cope with a crowded environment? Several alterations may be made in the shopping plan of the consumer. First, the shopper may attempt to reduce shopping time. The prevailing concern with time leads to a greater reliance on a shopping list and to a tendency to delay unnecessary purchases. (A typical comment from a respondent: "Your instincts are to get out of there as quickly as you can.") Interestingly, some shoppers appear to assign priorities to items, thus leaving some purchases for future shopping trips.

Another shopping adaptation strategy, consistent with the work of Milgram (1970), is to devote less time to each purchase decision. Obviously, a crowded store does not provide an atmosphere conducive to a careful evaluation of alternatives. Thus, a buyer relies on familiar brands and engages in minimal exploratory shopping. Similarly, the confusion and haste infused into the shopping trip by crowding may serve to reduce the amount of specific information processed by the consumer. The consumer's willingness and/or ability to examine unit price information, ingredient listings, nutritional data, and in-store advertising may be constrained by environmental conditions.

Crowding may likewise influence patterns of interpersonal communication within the store. Fewer special requests (e.g., meat cuts) may be made by shoppers. Shoppers engage in weaker forms of involvement with familiar persons, strangers, and store employees. The tempo and pace within the shopping environment may be a contributing factor to this reduced level of communicative involvement. Traffic patterns within a crowded store are analogous to those of a congested city. Established routes, directions, and speeds demand conformity.

central components of the model will be discussed in sequence. First, the factors that impinge on crowding are delineated. Second, coping or adaptation strategies initiated by shoppers in a crowded environment are explored. Third, possible behavioral and attitudinal outcomes stemming from a crowded environment are discussed.

Dimensions of Crowding

Crowding consists of two fundamental components: (1) physical density, and (2) perceived crowding. The physical density component involves the spatial limitations imposed on the buyer by the presence of other

## Outcomes of Crowding

What are the outcomes of shopping in a crowded environment? These comments from respondents offer some initial insights:

"Sometimes I'm confused by the purchases that I made ... Why did I buy margarine when butter was on sale."

"It takes a long time after the shopping trip to come out of the irritable feeling."

Several behavioral and attitudinal propositions that appear to be worthy of further exploration are presented in Figure 1. First, crowding influences the confidence of the shopper. In evaluating a shopping trip, the buyer may lack confidence that all necessary items were purchased and that the best values were secured. After shopping in a crowded store, consumers felt that they deviated from their shopping plan. Thus, the degree of satisfaction derived from the purchase selections may be affected by the conditions under which the shopper made those decisions. These conditions may likewise influence the shopper's image of the store.

Substantial research has been invested in studies of the factors that influence store image (Stephenson, 1969) (Berry, 1969). Environmental conditions, specifically physical density and crowding, may affect several dimensions of store image. A heavy concentration of shoppers may communicate a low price image more effectively than advertising. Additionally, the environmental condition encountered in the store may influence the shopper's perception of the store's personality (e.g., friendly-unfriendly) as well as several salient components of the product mix (e.g., assortment, quality, freshness).

How important are these considerations to initial store selection decisions? Clearly, a definitive answer cannot be derived from this preliminary analysis. A number of factors such as convenience, past experience, and shopping objectives, influence store selection decisions. However, a buyer particularly sensitive to the experiential state of crowding may engage in more extensive pre-planning in selecting a day and a time to shop at particular retail establishments.

### A Concluding Note

Crowding in retail shopping is an important environmental condition with implications for both the manager and the researcher. Interestingly, it presents a paradoxical problem for the manager. On the one hand, high density is required to maintain profitability; on the other hand, perceived crowding may have adverse effects on the shopper's attitudes and buying behavior. Under what conditions can density be at a maximum and perceived crowding at a minimum? Additionally, how can the retailer respond to shopper's adaptation strategies in mutually beneficial ways? Research into crowding will provide the knowledge to answer these important questions.

### References

- Leonard L. Berry, "The Components of Department Store Image: A Theoretical and Empirical Analysis," Journal of Retailing, 45(Spring, 1969), 3-20.
- J. B. Calhoun, "Population Density and Social Pathology," Scientific American, 1962, 206, 139-48.
- J. J. Christian, V. Flyger, and D. C. Davis, "Factors in the Mass Mortality of a Herd of Sika Deer," Cervus Nippon; Chesapeake Science, 1960, 1, 79-95.
- J. Desor, "The Psychology of Crowding: An Experimental Investigation," (Doctoral dissertation, Cornell University, 1970), Ann Arbor, Mich.: University Microfilms, No. 70-5763.
- J. A. Desor, "Toward a Psychological Theory of Crowding," Journal of Personality and Social Psychology, 1972, 21, 79-83.
- Rene Dubos, So Human An Animal (New York: Charles Scribner's Sons, 1968).
- Aristide Henri Esser, "A Biosocial Perspective on Crowding," in Joachim Wohlwill and Daniel Carson (eds.) Environment and the Social Sciences: Perspectives and Applications (Washington, D.C.: American Psychological Association, 1972, 15-28.
- R. Faris and H. W. Dunham, Mental Disorders in Urban Areas (Chicago: Phoenix Books, 1965 (originally published 1939)).
- Jonathan L. Freedman, Simon Klevansky and Paul R. Ehrlich, "The Effect of Crowding on Human Task Performance," Journal of Applied Social Psychology, 1971, 1, 7-25.
- Alfred E. Goldman, "The Group Depth Interview," Journal of Marketing, 26(July, 1962), 61-68.
- Philip Kotler, "Atmospherics as a Marketing Tool," Journal of Retailing, 49:4(Winter, 1974), 48-64.
- S. Milgram, "The Experience of Living in Cities," Science, 1970, 167, 1461-68.
- Harold Proshansky, William Ittelson, and Leanne Rivlin, "Freedom of Choice and Behavior in a Physical Setting," in Joachim Wohlwill and Daniel Carson (eds.) Environment and the Social Sciences: Perspectives and Applications (Washington, D.C.: American Psychological Association, 1972, 29-43.
- C. Schmid, "Urban Crime Areas," American Sociological Review, 25, 1960; 527, 542, 655-678.
- J. Smith, W. Form and G. Stone, "Local Intimacy in a Middle Sized City," American Journal of Sociology, 1954, 60, 276-85.
- P. Ronald Stephenson, "Identifying Determinants of Retail Patronage," Journal of Marketing, 33(July, 1969), 57-61.
- Daniel Stokols, "On the Distinction Between Density and Crowding: Some Implications for Future Research," Psychology Review, 1972, 79:3, 275-77.
- Daniel Stokols, Marilyn Rall, Berna Pinna, and John Schopler, "Physical, Social, and Personal Determinants of the Perception of Crowding," Environment and Behavior, 5(March, 1973), 87-115.
- Stuart Valins and Andrew Baum, "Residential Group Size, Social Interaction and Crowding," Environment and Behavior, 5(December, 1973), 421-39.
- Steven Zlutnick and Irwin Altman, "Crowding and Human Behavior," in Joachim Wohlwill and Daniel Carson (eds.) Environment and the Social Sciences: Perspectives and Applications (Washington, D.C.: American Psychological Association, 1972, 44-58.

## PROFILING THE MALE FASHION INNOVATOR--

### ANOTHER STEP

John J. Painter, University of Utah  
Kent L. Granzin, University of Utah

#### Abstract

A cluster technique isolated three groups of adult males on the basis of 116 attributes of 10 clothing items recently acquired. The groups were compared in terms of their demographics, AIO's, and media preferences and habits. Results extended previous studies to provide an enlarged profile of the fashion innovator.

#### Introduction

A number of recent studies have sought to identify and characterize the innovator (Baumgarten, 1975; Boone, 1970; Robertson, 1967; Robertson and Kennedy, 1968; Reynolds and Darden, 1972-73). While it would be desirable to generalize their findings beyond the product classes studied, it appears that in many cases such generalizations are unwarranted. Even within product classes noticeable differences appear to exist between men and women (Granzin and Painter, 1974; Painter and Pinegar, 1971).

This research concerns the important product category of men's clothing. Several studies of innovative behavior in the fashion area have generated profiles of the innovator that are enlightening, but incomplete. Granzin and Painter (1974) found demographic and media measures of the respondent useful for distinguishing innovators from non-innovators taken as a single group. Darden and Reynolds (1974) identified six consumer groups as based on innovative behavior. While they demonstrated the efficacy of demographic, AIO, and media variables for discriminating among the several groups, only one of the three product types used for group formation was clothing. Extension of their multiple group approach for the specific area of clothing purchases seems warranted.

The purpose of this study is to:

1. Provide further confirmation for the results of past studies in a multiple group context;
2. Consider new variables that may be useful in characterizing the male fashion innovator; and
3. Provide new, extended profiles of innovator and multiple non-innovator groups.

The results should provide marketers with insights that are useful in segmenting the market for men's clothing, whether their prime concern is the innovator or the non-innovator.

#### Method

A random cluster sample of adult males in the Salt Lake City area was interviewed in their homes by student interviewers. City blocks were selected on an equal-probability basis from 1970 census tract listings, and a constant percentage of homes in each block contacted. In the event of turn-downs, interviewers selected the next dwelling on their previously constructed block maps. The survey resulted in 271 completed responses to the self-administered questionnaire.

The criterion for characterizing the innovativeness of a respondent was based on an extensive inventory of clothing characteristics. The inventory listed attributes of the last piece of clothing he purchased or requested in 10 different clothing categories. By checking the appropriate set of attributes, the respondent indicated the nature of a particular clothing item. For example, he characterized a suit coat with respect to: weave (woven or double knit); style (double- or single-breasted); lapel width (narrow, medium, or wide); cut (straight or shaped); back (belted or non-belted); and color (solid, print/pattern, or stripes). Other clothing categories were: sportcoats, pants for wear with suitcoat/sportcoat, casual slacks, dress shirts, casual shirts, sweaters, ties, belts, and shoes. Purchases and requests for purchase by someone else were used because men receive at least one-fourth of certain clothing items as gifts (Ryan, 1966).

Cluster analysis of the 271 respondents on the 116 different 0-1 scores was "seeded" with perceptions of the managers of 25 local retail clothing stores. These managers responded to the same list of attributes of clothing as did the respondents by indicating those attributes they deemed most characteristic of items that were innovative in terms of recency of market introduction. Perceptions of the 25 managers were introduced along with the responses of the 271 respondents into a cluster routine developed by Veldman (1967) and based on the Ward (1963) criterion for hierarchical grouping. The original 296 "groups" were systematically combined on the basis of the similarity of the clothing attributes they checked. By the time only nine groups remained, all 25 managers were classified into the same group, attesting to the homogeneity of their perceptions. Clustering proceeded until only three groups remained. One group contained the 25 managers and 33 respondents. The managers were removed from the group and the remaining individuals were dubbed innovators. A three-group cutoff was chosen because of the large increase in the Ward error term in going from three to two groups and because three groups appeared to furnish an interesting contrast with usual innovator/non-innovator two-group comparison common in the literature of innovative processes. The 33 innovators constituted 12.2 per cent of the sample, a proportion comparable to the 12 to 15 per cent Wasson (1971) suggests make up the class of early adopters for products in general.

Predictor variables were of three types: AIO's, demographics, and media preferences and habits. The 18 AIO's used in this research were selected to represent shopping behavior, recreational preferences, and occupational dress needs. The latter two categories were included because one's dress was presumed related to his activities. Demographics represent time-tested predictors of innovative behavior, although their predictive value has sometimes been of secondary importance. Eighteen demographic variables were included.

Media preferences and habits were included because of the necessity for marketing decision-makers to communicate with their target market(s). The instrument

contained items relating to frequency and thoroughness of newspaper readership, number of magazines read regularly, and readership in different categories of magazines. Radio listening preferences referred to total hours a day of listening, preference for six classes of radio programming, and listening during eight periods distinguishable in a typical week. TV viewing related to total hours a day and preference for one of seven types of programming.

The three classes of predictor variables were analyzed separately by discriminant analyses on the three clothing groups. Use of the discriminant technique follows analyses of innovative behavior by Robertson and Kennedy (1968), Reynolds and Darden (1972-73), and Granzin and Painter (1974).

#### Results and Discussion

The cluster analysis generated three groups, which have been labeled A, B, and C. Table 1 represents the proportion of each group's members who purchased items possessing the listed attributes.

Group A (12.2 per cent of the sample) represents "innovators." They can be described as most likely to possess:

flare double-knit pants;  
double-knit, wide lapel, shaped cut, single-breasted and solid color coats;

print or pattern, no stripes, double-knit shirts;  
slip-on, non U-neck sweaters;  
no suede, high-heel or two-tone shoes;  
wide, no-stripe ties; and  
some wide belts.

Group B, the largest of the three groups (45.7 per cent of the sample), may be characterized as clothing "conservatives" who are most likely to own:

pants that are neither flare nor double-knit;  
coats that are not double-knit, have medium lapels, are not shaped and single-breasted;  
shirts that are conventional in style, not double-knit, not velour and without stripes;  
no turtleneck sweaters;  
no high-heel, suede or two-tone shoes; and  
no wide belts.

Group C (42.1 per cent of the sample) is more of a challenge to describe. Although some members of this group do try some new clothing items, their wardrobes do not contain as broad a selection of innovative items as those of the Innovators. They appear to be more selective. These "middlers" may be characterized by their ownership of:

flare pants;  
shaped, solid color, medium lapel coats;  
no velour shirts;  
some two-tone shoes;  
wide ties; and  
some wide belts.

TABLE 1

PROPORTIONS OF THREE GROUPS PURCHASING CLOTHING WITH PARTICULAR ATTRIBUTES

Clothing Attribute <sup>a</sup>	Group A (n=33)	Group B (n=124)	Group C (n=114)	Clothing Attribute <sup>a</sup>	Group A (n=33)	Group B (n=124)	Group C (n=114)
<u>Suit Coat</u>				<u>Casual Shirt</u>			
Single-breasted <sup>10</sup>	.70	.91	.51	Velour weave <sup>13</sup>	.21	.02	.05
Medium lapel <sup>36</sup>	.15	.75	.55	Conventional styled <sup>12</sup>	.18	.71	.46
Wide lapel <sup>3</sup>	.76	.05	.35	Print/pattern <sup>29</sup>	.39	.23	.40
Non-belted back <sup>19</sup>	.67	.83	.46	<u>Sweater</u>			
Striped pattern <sup>35</sup>	.15	.09	.19	Slip-on <sup>6</sup>	.55	.20	.69
<u>Sport Coat</u>				U-neck <sup>21</sup>	.06	.11	.23
Double-knit <sup>30</sup>	.70	.19	.40	Turtleneck <sup>37</sup>	.27	.03	.24
Double-breasted <sup>5</sup>	.06	.10	.36	<u>Tie</u>			
Wide lapel <sup>7</sup>	.76	.03	.29	Wide four-in-hand <sup>23</sup>	.91	.40	.59
Shaped cut <sup>2</sup>	.97	.15	.69	Striped <sup>20</sup>	.03	.40	.31
Solid color <sup>24</sup>	.52	.44	.60	"Wild" color <sup>31</sup>	.09	.00	.11
Belted back <sup>33</sup>	.36	.12	.46	<u>Belt</u>			
<u>Dress Pants</u>				Suede leather <sup>32</sup>	.06	.02	.16
Woven <sup>4</sup>	.06	.71	.32	Rope/Macrame <sup>27</sup>	.06	.01	.06
Flared legs <sup>15</sup>	1.00	.18	.85	Wide <sup>22</sup>	.36	.07	.39
No belt needed <sup>17</sup>	.33	.22	.18	<u>Shoes</u>			
Seam-opening pocket <sup>26</sup>	.15	.56	.41	Suede <sup>11</sup>	.00	.00	.10
<u>Casual Slacks</u>				Two-tone <sup>34</sup>	.42	.02	.37
Woven <sup>14</sup>	.09	.59	.13	High-heeled <sup>9</sup>	.36	.03	.13
Double-knit <sup>25</sup>	.48	.29	.39	<u>Dress Shirt</u>			
Flared legs <sup>1</sup>	.91	.15	.87	Double knit <sup>16</sup>	.48	.06	.21
Straight legs <sup>28</sup>	.09	.81	.10	Print/pattern <sup>8</sup>	.67	.14	.40
<u>Dress Shirt</u>				Striped <sup>18</sup>	.09	.10	.29

<sup>a</sup>Superscripts on clothing attributes denote order of entry into discriminant analysis.



Table 1 also indicates the results of the stepwise discriminant analysis of clothing attributes. The superscripts denote the order of entry of each attribute. Because the analysis operated on the same variables used in the group-forming cluster analysis, the presentation has no inferential value and is presented only to describe the clothing patterns characteristic of each group.

#### Demographics

Table 2 presents means and univariate F ratios on the 18 demographic variables. Ten of the 18 variables were significant at the .05 level. The demographic variables are listed in Table 2 in the order that they entered the stepwise discriminant analysis. The functions were highly significant, as attested by a Wilks' lambda of .550 and a multivariate F of 4.85 (d.f. = 36/502; p = .000). Age, education, and stage of life cycle were the best discriminators. Income, though significant on a univariate basis, proved redundant to more significant items and entered relatively late in the stepwise process. A separate univariate analysis of annual clothing expenditures also indicated no significant difference among the three groups with respect to amounts spent on clothing.

Group A (Innovators) is low in marital status, number of children, age, and life cycle stage, although scarcely different from Group C in these four items. Highest in education, the Innovator group is intermediate in income, as might be expected from a younger group yet to reach its potential earning power. Intermediate in apartment renting, highest in house renting, this group is low in house ownership. The mobility suggested by this finding is entirely consistent with their younger age.

Group B (Conservatives) presents the reverse image of Group A. Highest in marital status, number of child-

ren, age, and life cycle stage--all by a wide margin--this group is simply older and more established in a life pattern than the other two groups. Lowest in education, this Conservative group has nonetheless achieved the highest income level. With a proportion of home ownership almost twice that of the next group, Group B can be portrayed as settled and having the least mobility.

Group C (Middlers) presents almost the same youthful pattern as Group A, but is lower on education level and income. Its higher proportion of apartment rental and lower inhabitation of a house, either through rental or ownership, suggests an even greater mobility than that seen for Group A. Group C apparently falls lower in the social order than Group A, and over time may be expected to converge upon the purchasing patterns of Group B to a greater extent.

#### Activities, Interests and Opinions

Table 3 presents means and univariate F ratios on 18 AIO's. Ten of the 18 variables were significant at the .05 level. Table 3 also indicates the results of the stepwise analysis of the AIO's, with each variable listed in the order of its entry. Again, the 18-variable discriminant functions were highly significant, with a Wilks' lambda of .635 and a multivariate F of 3.56 (d.f. = 36/502; p = .000). Purchase of new items before friends, enjoyment of active sports, and the purchase of name brand clothing to insure quality were the most important discriminators.

Group A (Innovators) indicates a relatively strong interest in trying new items on the market. This interest reflects more than curiosity, given the group's desire to look nice to women. Perhaps importantly, the group does not express a commensurately significant desire to dress well. One way of looking nice may be to buy name brand clothing. The Innovative group does

TABLE 2  
MEANS AND UNIVARIATE F RATIOS FOR 3 GROUPS  
ON 18 DEMOGRAPHICS

Demographics <sup>a</sup>	Group A	Group B	Group C	F Ratio
Age (years)	27.91	46.08	28.90	69.72**
Education (years of schooling)	16.61	13.94	14.24	5.36*
Life cycle stage (1 to 7)	2.18	4.22	2.18	66.03**
Rents apartment (0=no; 1=yes)	0.21	0.10	0.38	14.54**
Rural background (0=urban; 1=rural)	0.18	0.31	0.35	1.70
Owns house (0=no; 1=yes)	0.42	0.81	0.39	29.63**
Number of clubs, associations, organizations	1.45	1.09	1.19	0.99
Neighborhood (1 to 9 quality scale)	4.67	4.90	5.11	1.29
Number of children	1.03	2.73	1.07	29.17**
Oriental race (0=no; 1=yes)	0.00	0.01	0.01	0.14
Caucasian race (0=no; 1=yes)	1.00	0.94	0.94	1.03
Marital status (0=single; 1=married)	0.61	0.92	0.55	25.28**
Income (1 to 7)	4.73	5.42	4.43	8.39**
Occupation (1963 NORC prestige scale)	70.97	69.57	68.91	0.39
Other dwelling (0=no; 1=yes)	0.09	0.01	0.04	3.17*
Other race (0=no; 1=yes)	0.00	0.02	0.04	0.65
Negroid race (0=no; 1=yes)	0.00	0.02	0.03	0.52
Rents house (0=no; 1=yes)	0.27	0.08	0.19	5.22*

\*p<.05

\*\*p<.001

<sup>a</sup>Variables are listed in the order in which they entered the stepwise discriminant analysis.

not differ from the other groups in regard to the greater importance of quality and/or service over price. This group is least prone to dislike shopping for clothes, although it is not the most likely to know what it wants when going shopping. Perhaps clothes shopping is a pleasant experience for Group A. More interested in socializing, the group enjoys outdoor sports, quite possibly of a competitive nature. Because these recreation activities are relatively active, the group may generate more perspiration than accumulate actual dirt.

Group B (Conservatives) in most cases lies at the opposite end of the continua from Group A. Showing a relative disinterest in new items and in looking nice to women, this group lies intermediate on buying name brand clothing. The group is most likely to know what it wants when shopping, but this finding may be a reflection of its avoidance of new and different items. It is also highest on disliking the clothes shopping process. Least sociable, Group B is least likely to favor active and outdoor sports.

Group C (Middleers) presents, in most cases, a compromise between the two other groups. Its interest in new items, however, is high. Looking nice to women is a concern, but one not reflected in purchase of name brand clothing. Intermediate on dislike for clothes shopping, Group C is least likely to know what it wants when shopping. Almost as sociable as Group A, Group C is highest on active sports, relatively high on preference for outdoor recreation, and most likely to get dirty during these activities.

#### Media Preferences and Habits

Table 4 contains the means, proportions, and univariate F ratios for media habits and preferences. Of the 39

media variables, only nine had univariate F ratios significant at the .05 level. The 39-variable discriminant analysis gave a Wilks' lambda of .566 and a multivariate F of 1.939 (d.f. = 78/460; p = .000). Table 4 also presents the results of the stepwise analysis. Only thirteen of the variables entered here. Of the first five variables entering, two were related to radio (rock music and contemporary popular music), two were magazines (Reader's Digest and Playboy), and the other was the average amount of time spent watching television.

Time patterns of radio listening differ surprisingly little among the three groups, with the only exception being the 7 p.m. weekend evening time slot. However, Group A (Innovators) is high on listening to stations that play rock music and low on those featuring traditional (i.e., generally restrained) popular music. Lowest on interest in watching television, the Innovators show the highest desire to watch sports programming, which usually features competitive athletic contests. Most thorough in newspaper readership, Group A also leads in reading magazines featuring spectator sports and athletic events. The latter finding furnishes confirmation of its television preferences.

Group B (Conservatives) stands lowest in weekend evening radio listening, while its general listening pattern shows a clear preference for traditional popular music stations and an avoidance of rock stations. The group is highest in TV viewing, intermediate in sports interest, high in thoroughness of reading the newspaper, and lowest in readership of magazines featuring spectator sports and athletic events. Perhaps its interest in TV sports viewing is more compatible with a more sedentary life style than an active interest in competitive sports. The group's strong interest in Reader's Digest

TABLE 3  
MEANS AND UNIVARIATE F RATIOS FOR 3 GROUPS  
ON 18 AIO'S

AIO <sup>a,b</sup>	Group A	Group B	Group C	F Ratio
Often try new items before friends and neighbors	44.2	69.7	49.8	23.50**
Enjoy active sports	21.7	39.1	20.9	15.53**
Buy name brand clothing to insure quality	34.6	43.9	53.2	6.31*
Like to try new and different products	41.3	54.9	37.9	14.30**
Shop for clothing specials	55.1	53.8	59.4	1.10
Like to get out evenings and socialize	30.8	48.1	32.1	12.78**
My favorite recreation gets me dirty	56.6	57.1	46.1	4.30*
One can save money through shopping for clothing bargains	47.2	40.3	36.1	2.21
It is important to look nice to women	27.2	43.8	30.4	9.91**
Know what I want when shopping for clothes	40.5	31.9	41.4	3.45*
Watch ads for clothing sales	51.6	57.4	59.1	0.82
It's important to dress well	33.1	42.3	41.6	1.53
Dislike shopping for clothes	59.3	45.6	53.0	3.13*
Getting quality/service when buying clothes is more important than price	30.2	35.4	26.5	0.63
Favor outdoor recreation activities	31.2	40.4	33.4	4.11*
My usual job gets me dirty	73.7	64.9	61.1	1.81
Often buy a new item just to see what it's like	66.6	76.6	72.0	2.41
Choice of clothes influenced by employer	67.0	65.5	61.6	0.64

\*p<.05

\*\*p<.001

<sup>a</sup>Low score denotes agreement with statement.

<sup>b</sup>Variables are listed in the order in which they entered the stepwise discriminant analysis.

and clearly lowest readership of Playboy are consistent with the older and more conservative pattern emerging from this analysis.

Group C (Middlers) is highest on weekend-evening radio listening and shows the highest preference for rock music. Low in TV viewing, the group is also lowest in watching sports broadcasts. Least thorough in reading newspapers, Group C lags in Reader's Digest and maga-

zines devoted to competitive sports. However, the group stands highest in Playboy readership.

#### Summary and Conclusions

The results of this study reveal profiles of the clothing innovator and non-innovator that have a good deal of similarity--and some dissimilarity--with the profiles generated by previous studies.

TABLE 4  
MEANS, PROPORTIONS, AND UNIVARIATE F RATIOS  
FOR 3 GROUPS ON 39 MEDIA ITEMS

Media Item <sup>a,b</sup>	Group A	Group B	Group C	F Ratio
<u>Radio</u>				
Average hours per day	2.15	2.07	2.68	2.31
Listen 7-9 am Mon-Fri	0.61	0.60	0.61	0.04
Listen 9-4 Mon-Fri	0.18	0.21	0.21	0.07
Listen 4-7 Mon-Fri	0.39	0.30	0.32	0.54
Listen 7 pm Mon-Fri	0.21	0.27	0.31	0.59
Listen 7-9 am Sat-Sun <sup>7</sup>	0.06	0.18	0.15	1.39
Listen 9-4 Sat-Sun	0.27	0.23	0.22	0.21
Listen 4-7 Sat-Sun	0.15	0.09	0.18	2.00
Listen 7 pm Sat-Sun	0.21	0.15	0.28	3.33*
Traditional popular music	0.18	0.40	0.18	9.02**
Rock music <sup>1</sup>	0.39	0.12	0.49	22.73**
Country music	0.03	0.10	0.06	1.06
Contemporary popular music <sup>3</sup>	0.30	0.21	0.20	0.81
Talk shows	0.03	0.06	0.03	0.74
Miscellaneous; no preference	0.06	0.11	0.04	2.06
<u>Television</u>				
Average hours per day <sup>4</sup>	1.85	2.34	1.88	3.47*
Prefer sports	0.61	0.45	0.35	3.72*
Prefer westerns	0.06	0.12	0.11	0.50
Prefer drama	0.03	0.05	0.03	0.43
Prefer variety	0.06	0.09	0.07	0.22
Prefer documentary, fact <sup>8</sup>	0.09	0.13	0.23	2.90
Prefer comedy	0.06	0.08	0.13	1.18
Prefer mystery	0.12	0.13	0.10	0.32
<u>Newspaper</u>				
Frequency of readership (4 point scale)	2.27	2.27	2.31	0.01
Thoroughness of readership (4 point scale) <sup>12</sup>	2.60	2.57	2.30	3.26*
<u>Magazines</u>				
Number of magazines <sup>13</sup>	1.39	1.64	1.46	1.36
Spectator sports and athletic events <sup>6</sup>	0.27	0.09	0.13	4.00*
Hunting, fishing, field sports	0.06	0.06	0.04	0.48
Electronics, mechanics	0.03	0.06	0.06	0.20
Geography <sup>11</sup>	0.00	0.10	0.09	1.86
Pictorial current interest	0.24	0.19	0.19	0.25
Reader's Digest <sup>2</sup>	0.15	0.26	0.07	7.54**
Newsweeklies	0.52	0.38	0.54	2.09
Business and commerce	0.03	0.09	0.06	0.71
Literature and theatre	0.09	0.06	0.06	0.12
Playboy <sup>5</sup>	0.18	0.07	0.25	7.60**
Religious	0.03	0.14	0.06	1.89
Trade publications <sup>9</sup>	0.06	0.12	0.03	2.41
Miscellaneous low circulation <sup>10</sup>	0.06	0.19	0.16	1.00

\*p<.05

\*\*p<.001

<sup>a</sup>Superscripts on media items denote order of entry into discriminant analysis.

<sup>b</sup>Weight of readership, rather than proportion of respondents reading magazines in a given category, is given as ratio of number of magazines in a given category read by group as a whole to total number of respondents in group.

In general, this three-group analysis of the clothing market seems to present two definite extremes, as portrayed by Groups A and B. This study provides additional evidence that the innovator is most likely to be young, have few children, and be mobile. At the other extreme, it portrays the clothing conservative as least prone to innovation in general, much older, having a larger family, and not likely to be mobile. Doubt remains regarding the education level of the innovator because this study suggests the innovator has a higher education level than the non-innovator. On the other hand, the Darden and Reynolds study (1974) found the innovator to have a lower education level.

It also appears that clothing innovation is not importantly influenced by income level or clothing expenditures; rather, the issue is the way income is allocated to particular clothing purchases. This result is consistent with the finding of Darden and Reynolds (1974).

The profiles of the innovator and non-innovator are also expanded by use of AIO's. They characterize the innovator as more interested in brand name clothing and quality than in price. His desire to socialize is consistent with the results of a recent study of the "innovative communicator" by Baumgarten (1975). The innovator's strong predisposition to buy and try new items before friends and neighbors suggests that his innovative tendencies may extend beyond the clothing product category.

Differences in media habits and preferences go beyond magazine readership. For radio, the programming itself seems to be relevant. In particular, the rock music format seems to be an important variable, as it was in the Baumgarten study (1975). While the innovator reads the newspaper more thoroughly, he has a tendency to spend less time viewing television. He is more interested in sports programming than the non-innovator, and this interest in TV sports is consistent with his high level of interest in active sports participation. In general, the magazine readership results are consistent with findings of other studies (Darden and Reynolds, 1974; Granzin and Painter, 1974).

Given the results of this study, the innovator is fairly well profiled with respect to the simpler types of characterizing variables. In their future work researchers would do well to investigate more detailed items such as the form of the consumer's clothing search process and the way in which he decides on stores, brands, and clothing items.

Also, the usefulness of separating non-innovators into more than one group is questionable, given the static nature of the criterion measures. What is needed is a means for developing purchase groupings with more information content for the marketer. A methodological shift to longitudinal study of the patterns of clothing purchase over time could bring fresh insights. By recording the post-introduction timing of consumers' purchases, the time pattern of their adoption cycles could be obtained, and these patterns aggregated to provide more meaningful groupings of consumers than those obtained by present approaches. Such classification of purchase groups by the nature of their adoption patterns would be more useful to marketers seeking to formulate strategies for various stages of a clothing item's product life cycle.

#### References

Steven A. Baumgarten, "The Innovative Communicator in the Diffusion Process," Journal of Marketing Research,

12(February, 1975),12-18.

Louis E. Boone, "The Search For the Consumer Innovator," Journal of Business, 43(April, 1970),134-140.

William R. Darden and Fred D. Reynolds, "Backward Profiling of Male Innovators," Journal of Marketing Research, 11(February, 1974),79-85.

Kent L. Granzin and John J. Painter, "Characteristics of Male Fashion Innovators: A Discriminant Analysis," Mississippi Valley Journal of Business and Economics, 10(Fall, 1974),59-66.

John J. Painter and Max L. Pinegar, "Post-high Teens and Fashion Innovation," Journal of Marketing Research, 8(August, 1971),368-69.

Fred D. Reynolds and William R. Darden, "An Analysis of Selected Factors Associated with the Adoption of New Products," Mississippi Valley Journal of Business and Economics, 8(Winter, 1972-73),31-42.

Thomas S. Robertson, "Determinants of Innovative Behavior," Proceedings, Winter Conference, American Marketing Association, 1967, 328-32.

Thomas S. Robertson and James N. Kennedy, "Prediction of Consumer Innovators: Application of Multiple Discriminant Analysis," Journal of Marketing Research, 5(February, 1968),64-69.

Mary S. Ryan, Clothing: A Study in Human Behavior (New York: Holt, Rinehart and Winston, 1966).

Donald J. Veldman, Fortran Programming for the Behavioral Sciences (New York: Holt, Rinehart and Winston, 1967).

Joe H. Ward, Jr., "Hierarchical Grouping to Maximize an Objective Function," Journal of the American Statistical Association, 58(March, 1963),236-44.

Chester R. Wasson, Product Management: Product Life Cycles and Competitive Marketing Strategy (St. Charles, Illinois: Challenge Books, 1971).

# FASHION INVOLVEMENT AND BUYING BEHAVIOR:

## A METHODOLOGICAL STUDY

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

### Abstract

This paper reports on continuing exploratory development of a comprehensive measure of "fashion involvement". The research presented here is the logical extension of a ten year tradition in fashion segmentation research and the fourth in a series of papers reporting on a major fashion research program in Canada. The "index of fashion involvement" developed here is validated with an independent AIO measurement methodology and is utilized to demonstrate that the highly fashion involved consumer is also the heavy clothing fashion buyer.

### Introduction

The process by which new clothing and apparel concepts, "style statements", and tastes continually cycle across the population has been the subject of popular commentary for centuries. Aesthetic expression in dress and adornment is the classic statement of fashion, and has been the focus of a variety of strong methodological research papers over the past fifteen years.<sup>1</sup>

Contemporary fashion research indicates that clothing consumers are distributed across a broad spectrum of general fashion involvement or fashion consciousness (Sproles & King, 1973). Conceptually, this research has been developed and operationalized within the theoretical and methodological framework of the adoption and diffusion of innovations research tradition. Within that conceptual framework, contemporary fashion research has been, for the most part, focused on two key market segments: the fashion innovators or early adopters of new clothing concepts or "style statements"; and the fashion opinion leaders or interpersonal communicators, the local or peer group legitimizers of new clothing concepts.

The research presented here continues and builds upon those earlier efforts. Specifically, the paper is three-dimensional in focus:

1. The research further probes the concept of general fashion involvement. A "fashion involvement" index is developed and applied to similar male and female populations.
2. The issue of "index validation" is examined through the utilization of a comparative "fashion involvement" measurement by means of fashion specific life-style/AIO factor analytic techniques.
3. The concept of fashion market segmentation based on degree "fashion involvement" is probed with specific attention to unit and dollar volume fashion consumption.

<sup>1</sup>For example see King (1963), Summers (1970), Baumgarten (1975), Grindereing, (1967), and Reynolds and Darden (1972).

### Contemporary Fashion Theory

#### Historical Measurement

Historically, contemporary fashion research has focused on the fashion innovator and the fashion opinion leader as prime sales targets and key links to the volume fashion market. Sophisticated methodological studies have been painstakingly designed to define, isolate, and profile these segments in a structured "block-by-block" manner over the past fifteen years, beginning with the research of King (1963).

In each of these cases, a variety of research measures have been utilized to develop an index (of innovativeness, opinion leadership, etc.) upon which the population is continuously distributed. For example, King probed the traditional fashion adoption process model known as the "trickle down" theory by conducting an exploratory study of fashion change agents in the fashion adoption process within the context of women's millinery. The crucial variable in that Boston survey was the inclusion of the time of adoption in the fashion season in the research design.

Later, Summers (1970), focused on identifying and profiling opinion leaders/interpersonal communicators in the fashion adoption process. In this research, a sophisticated opinion leadership indexing technique was developed and tested. The basic conclusion in this and earlier research was that innovators and opinion leaders played key roles in directing fashion adoption and represented discrete market segments within social strata.

Carrying this structure a step further, Baumgarten (1975) researched the intersection of the fashion innovator and the fashion opinion leader: the "innovative communicator". Theoretically, the innovative communicator performs jointly and simultaneously in the early buyer and new style communication and influential roles.

At the same time, King, Summers, Baumgarten, and others had been monitoring "fashion awareness" across a number of populations over several years. In the United States and Canada, as reported by King, Tigert and Ring (1975) a comparative fashion awareness data base spanning thirteen years has now been established.

The research described briefly here, and other related efforts indicate that "fashion innovativeness", "fashion opinion leadership", "fashion awareness", etc., exist as related but distinct continuums among the consumer population. Sproles and King (1973) suggest that such "fashion continuums" can be defined across a wide range of fashion-related activities and behavioral dimensions. What seems clear is that the population is differentially involved in fashion along a number of dimensions.

#### The Concept of Fashion Involvement

Given a number of empirically defined, distinct fashion behavioral dimensions, what is the joint and simultaneous interaction effect of that intersection? Baumgarten (1975) explored this question in part, by researching

the intersection of the fashion innovator and the fashion opinion leader, the innovative communicator. In that research, the focus was on two dimensions of fashion related activity and behavior. It logically follows that there are several other additional dimensions which should also be included in this kind of analysis.

The concept of "fashion involvement" is based essentially on three propositions:

1. The population is distributed along a broad continuum in terms of fashion behavioral activities;
2. The population is also distributed on a unidimensional continuum for each of these fashion behavioral activities;
3. For several specific fashion behavioral activities, these continuums have been and can be researched and identified for specific geographic submarkets.

Theoretically, an overall fashion involvement continuum can be defined based on the aggregate effect of a variety of important fashion behavioral activities. Sproules and King (1973) suggest, based on previous fashion segmentation research, that there are at least five important dimensions of the aggregate fashion involvement continuum.

They are:

1. Fashion innovativeness and time of purchase. The continuum which ranges from the early adopting and experimenting consumer to the late buying, conservative consumer;
2. Fashion interpersonal communication. A continuous dimension which describes the relative communicative and influential power of the consuming population at conveying fashion information;
3. Fashion interest. A continuum ranging relatively from the highly interested fashion consumer to the totally non-interested buyer;
4. Fashion knowledgeability. Consumers range from those who are relatively knowledgeable about fashions, styles and trends to those who have no insight into the fashion arena;
5. Fashion awareness, and reaction to changing fashion trends. A continuum ranging from the consumer who is very actively monitoring the style trends to the totally non-aware individual.

As the first step toward building an aggregate index of fashion involvement, these five selected dimensions were utilized for initial analysis purposes. For each of the selected fashion behavioral dimensions, there exists a history of empirical research and tested measurement technology. Based on that research,<sup>3</sup> five individual questions, each a field tested measure of one of the five basic fashion behavioral dimensions, were selected for the development of "fashion involvement index".

<sup>3</sup>These measures have been developed and tested over the course of four major consumer research programs supervised by C.W. King. For the most recent work on the individual measures, see King, Tigert and Ring (1975) where the questions were tested individually preparatory to their aggregations into an index.

Those questions are:

1. Fashion innovativeness and time of purchase. In general, would you say you buy new men's clothing fashions earlier in the season, about the same time, or later in the season than most other men?

Earlier in the season than most other men.....

About the same time as most other men.....

Later in the season than most other men.....

2. Fashion interpersonal communication. Would you say you give very little information, an average amount of information, or a great deal of information about new men's clothing fashions to your friends?

I give very little information to my friends.....

I give an average amount of information to my friends...

I give a great deal of information to my friends.....

3. Fashion interest. In general, would you say you are less interested, about as interested, or more interested in men's clothing fashions than most other men?

Less interested than most other men.....

About as interested as most other men.....

More interested than most other men.....

4. Fashion knowledgeability. Compared with most other men, are you less likely, about as likely, or more likely to be asked for advice about new men's clothing fashions?

Less likely to be asked than most other men.....

About as likely to be asked as most other men.....

More likely to be asked than most other men.....

5. Fashion awareness, and reaction to changing fashion trends. Which one of the statements below best describes your reaction to changing fashions in men's clothes? (Even though there may be no statement listed which exactly describes how you feel, make the best choice you can from the answers listed).

I read the fashion news regularly and try to keep my wardrobe up-to-date with the fashion trends.....

I keep up-to-date on all the fashion changes although I don't always attempt to dress according to those changes.....

I check to see what is currently fashionable only when I need to buy some new clothes.....

I don't pay much attention to fashion trends unless a major change takes place.....

I am not at all interested in fashion trends.....

Simple sum scores across these five questions were computed for each consumer, resulting in a thirteen point continuum upon which the population is distributed.

The fashion awareness question has a five point response scale, compared to a three point scale for the remaining four dimensions. Therefore, the fashion awareness question has a proportionally higher weighting in the overall sum score. Further development using factor scores, standardized scores, and additional fashion dimensions is in progress.

This analysis is based on data gathered through the Toronto Retail Fashion Market Segmentation Research Program.<sup>4</sup> This program developed from a collaboration by King, Tigert and Ring in the fashion arena, beginning with a 1974 study which integrated fashion segmentation dimensions with retail patronage analysis. From that initial effort a series of research projects evolved into an integrated research program. The broad objective of this program is to gain a better understanding of the dynamics of the fashion retailing process through the integration of contemporary fashion segmentation, lifestyle/AIO, and retail image methodologies within the broader program. This paper focusses on only one dimension of that process.

#### Sampling

The data base for this research was collected in March/April 1975 as part of the Toronto Retail Fashion Market Segmentation Research Program. The sample was part of a specially recruited panel of husband and wife pairs. The data were gathered by mailed self-administered questionnaires from an area probability sample in Census Metropolitan Toronto, Canada. A total of approximately 1000 husband and wife pairs completed and returned the questionnaires.

#### Sample Distributions

The objective of this section is to graphically illustrate the aggregate male and female fashion involvement sample distributions as determined by the fashion involvement index.

For each respondent, a simple sum score was calculated over the five selected questions listed earlier. Using a scale running from 5 (low) to 17 (high), the females and males were independently measured and analyzed. The distributions over the approximately 1000 respondents for both the male and female samples are displayed in Figure 1. The female distribution is uni-modal, and the male distribution is slightly bimodal. However, significantly different mean scores were obtained for the two samples.<sup>5</sup> The female sample obtained a mean score of 9.7, while the male sample yielded a mean score of 8.4.

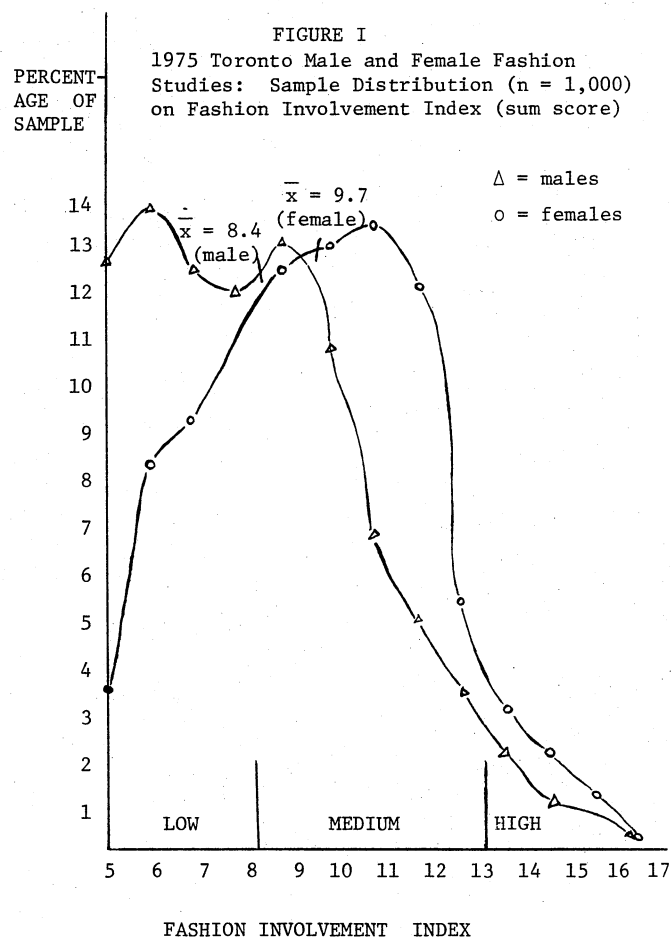
Distributionally, two observations are noteworthy.

1. Both the male and female distributions are strongly skewed toward the non-fashion involved end of the continuum. This may indicate that, in general, Toronto consumers are not highly fashion involved.
2. Toronto females are more "involved" in fashion than are Toronto males, based on mean sample differences.

<sup>4</sup>For a more complete description of the Toronto Retail Fashion Market Segmentation Research Program, see King, C.W., and L.J. Ring, "Retail Fashion Segmentation Research: Development and Implementation", in Bent Stidsen, editor, *Marketing in the 1970's and Beyond*, Canadian Association of Administrative Sciences (Proceedings), Edmonton, Canada, 1975.

<sup>5</sup>The two sample means were five average standard errors apart. The statistical test for the hypothesis of no difference was rejected at the .001 level.

For this research, an attempt was made to validate the concept of fashion involvement, and this fashion involvement index by measuring it in a second and entirely



independent and unrelated manner to what has historically been the case in fashion segmentation research. Factor analytic techniques have been utilized across a variety of fashion specific lifestyle/AIO questions to obtain a "fashion involvement factor".

Over the past several years, the authors have conducted numerous in-depth, focused group discussions with men and women about fashion in both the United States and Canada. Based on that probing, a number of hypotheses have been generated about fashion attitudes and activities and fashion buying behavior which carry strong implications for fashion retailers.

Those hypotheses have been stated in 24 different lifestyle measures. The measures ranged across social attitudes (non-fashion specific) about unionism, sexual behavior, religion, and deviant behavior. Additionally, the measures covered a wide range of fashion related dimensions, such as commitment to fashionability, fashion trends, and clothes shopping. The measures have been operationalized around a six point (strongly agree to strongly disagree) Likert Scale. The questions were included in the questionnaire utilized in data collection for this paper. In addition, many of these same questions have been used in previous fashion related

research efforts.

The responses to the 24 lifestyle/AIO questions were submitted to a principal factor analysis and varimax rotation. The first and strongest factor in both the male and female data is employed here as another measure of fashion involvement/consciousness and has been so labeled. The rotated factor loadings are reported in Table I for both the male and the female data. A similar factor has been produced and reproduced in several previous studies in both the United States and Canada (Wells & Tigert, 1971; and Tigert, 1973). This factor has been both highly reliable and stable regardless of the number of mix of lifestyle questions analyzed in any particular study. Additionally, the factor has been utilized as a predictor of fashion buying behavior, fashion magazine readership and other fashion attitudes.

TABLE I

Toronto, 1975:  
Fashion Involvement/Consciousness Factor

Variables	Loadings	
	Male	Female
1. I usually have one or more outfits of the very latest style.	.66	.60
2. An important part of my life and activities is dressing smartly.	.64	.54
3. I like to shop for clothes.	.49	.44
4. I like to think I'm a bit of a swinger.	.49	.47
5. For my fashion needs, I am increasingly shopping at boutiques or fashion specialty stores rather than department stores.	.47	.32
6. When I must choose between the two I usually dress for fashion, not comfort.	.39	.39

The Fashion Involvement Index vs. The Fashion Involvement Factor

Distributionally, both the fashion involvement index and the fashion involvement factor obtained higher mean scores for females than males. For the factor, females obtained a sample mean which was 6.5 percent higher than the male sample mean score (rejecting the no difference hypothesis of the .001 level). For the "fashion involvement index", the female mean was 10 percent higher than the male mean. In addition, the shapes of both sets of distributions were highly similar.

As a measure of that similarity, correlation analysis was run between the index and factor sum scores. The two measures obtained a correlation coefficient of .633 in the male case, and .566 in the female example.

Therefore, the conclusion is that while the index and the factor do not measure exactly the same thing, they are similar enough to suggest that the concept of "fashion involvement" is in fact a valid one.

Fashion Involvement and Buying Behavior

To date, contemporary fashion researchers have devoted their primary research efforts toward the identification and profiling of key market segments thought to be driving or influencing the fashion adoption process. Innovators and opinion leaders have been profiled by several fashion research teams across a wide variety of sociological, psychological, lifestyle, demographic, and other characteristics.

What has not been explored in great depth, however, is the actual buying and purchasing behavior of highly fashion involved consumers. In a paper presented before the American Marketing Association in August, 1975, the authors probed retail clothing fashion store shopping behavior across several dimensions of fashion involvement. A methodology was developed for the application of contemporary fashion segmentation technology to a specific fashion retailing context.

In that Toronto-based research, it was clear that significant differences existed in the levels of fashion involvement among the patrons of major Toronto men's wear chains. For example, in general the fashion specialty chains had a much more "fashion involved" customer mix than did the mass merchandiser set of stores.

In order to further probe consumer fashion purchasing behavior, this section of the paper examines the relationship between fashion involvement and unit and dollar volume fashion buying behavior.

Toronto Male Clothes Buying

As a part of the 1975 Toronto Male Fashion Study which was described in the data base section of this paper, respondents were asked to estimate the number of units they purchased over the past year across 11 fashion product/object categories. In addition, respondents estimated the average price they paid across those same 11 categories.

For this paper, the 11 categories were collapsed into three major classes, "big ticket items", "dressy accessories", and "casual wear". Three particular fashion product/objects were selected, one from each class, as representative of Toronto male fashion buying in general.

The analysis here focuses on suits, dress shirts, and casual slacks (defined as "jeans, corduroy, denim"). For each of these three categories, cross classification analysis was performed across unit and dollar volume purchased against "fashion involvement" as measured by the index described earlier in this paper.

As an operating hypothesis, the expectation was that the more highly fashion involved consumers would be heavier clothing buyers in terms of both price (price per unit bought) and volume (number of units bought) than less fashion involved individuals. That hypothesis was, in fact, supported. Those results are depicted in Tables 2, 3, 4.

Cross Classification

For purposes of the cross classification analyses, the fashion involvement distribution was arbitrarily segmented into "low, medium, and high" fashion involved groupings. Those groupings are reflected in Figure 1 which was presented earlier.

The cross classification analysis sequence graphically illustrates some rather dramatic differences in fashion purchasing behavior between the "high" fashion involved group and the "medium to low" groups.



For example, turning first to suits as a fashion product/object category (Table 2), the "high" fashion involved consumers are clearly the heavy buyers in terms of both numbers of suits and in terms of dollars per suit. Twenty percent of "high" fashion involved people said they bought four or more suits in the past year. This is four times the number in either the "medium" or "low" fashion involved groups. In terms of prices, 47 percent of "high" fashion involved men said they paid over \$200 per suit on average, while the "medium" and "low" fashion involved groups had less than 20 percent purchasing suits over \$200. For both unit and dollar analyses, the differences between the groups were statistically significant based on a chisquare test at the .001 level.

TABLE 2

Fashion Involvement and Suit Buying

1. Units

Number of Suits Purchased in Past Year	Level of Fashion Involvement <sup>1</sup>			Base
	Low	Medium	High	
0 or 1	82.0%*	63.1%	27.3%	(720)
2 or 3	17.4	30.2	52.2	(244)
4 or more	0.6	6.7	20.5	( 40)
Total	100.0%	100.0%	100.0%	
Base	(540)	(420)	( 44)	

<sup>1</sup>The differences were statistically significant at the .001 level based on X<sup>2</sup> analysis.

\* READ: 82.0 percent of the "low" fashion involved group bought one or no suits during the past year, while 63.1 percent of the "medium" group and only 27.3 percent of the "high" group bought one or no suits.

2. Prices

Average Price Paid per Suit Last Year	Level of Fashion Involvement <sup>1</sup>			Base
	Low	Medium	High	
Less than \$100	28.6%*	15.4%	5.3%	(125)
\$100 - \$199	54.3	66.3	47.3	(364)
\$200 and over	17.1	18.3	47.4	(119)
Total	100.0%	100.0%	100.0%	
Base	(265)	(305)	( 38)	

<sup>1</sup>The differences were statistically significant at the .001 level based on X<sup>2</sup> analysis.

\* READ: 28.6 percent of the "low" fashion involved group said they paid an average of less than \$400 per suit during the past year. 15.4 percent of the "medium" group and 5.3 percent of the "high" group paid less than \$100.

With regard to the dress shirt analysis, Table 3, again the "high" fashion involved consumers are the heavy buyers. 51.1 percent of the "high" fashion involved group said they bought six or more dress shirts over the year. In contrast, only 29.6 percent of the "medium" group and

13.7 percent of the "low" group said they purchased six or more dress shirts during the same period. Price-wise the differences are similar, with well over fifty percent of "high" fashion involved people paying the high prices, over twice as many as the "medium" and "low" groups.

In the casual slacks analysis sequence, Table 4, the differences were not as dramatic as in the suit and dress shirt analyses. However, the "high" fashion involved group still bought more pairs of casual slacks, and paid more money per pair of casual slacks than did the members of the "medium" and "low" fashion involved groups. The smaller differences here may be attributed to the fact that casual slacks are generally not considered as much of a "fashion" apparel item as suits and dressy shirts.

TABLE 3

Fashion Involvement and Dress Shirt Buying

1. Units

Number of Dress Shirts Purchased in Past Year	Level of Fashion Involvement <sup>1</sup>			Base
	Low	Medium	High	
2 or less	55.0%*	32.3%	20.9%	(442)
3, 4 or 5	31.3	38.1	28.0	(341)
6 or more	13.7	29.6	51.1	(221)
Total	100.0%	100.0%	100.0%	
Base	(540)	(421)	( 43)	

<sup>1</sup>The differences were statistically significant at the .001 level based on X<sup>2</sup> analysis.

\* READ: 55.0 percent of the "low" fashion involved group said they purchased 2 or less dress shirts during the past year.

2. Prices

Average Price Paid per Dress Shirt Last Year	Level of Fashion Involvement <sup>1</sup>			Base
	Low	Medium	High	
Less than \$9.00	38.3%*	24.8%	9.0%	(255)
\$9.00 - \$14.00	38.9	42.2	31.8	(333)
\$15.00 and over	22.8	33.0	59.2	(244)
Total	100.0%	100.0%	100.0%	
Base	(412)	(376)	( 44)	

<sup>1</sup>The differences were statistically significant at the .001 level based on X<sup>2</sup> analysis.

\* READ: 38.3 percent of the "low" fashion involved group said they paid an average of less than \$9.00 per dress shirt during the past year.

Overall, the cross classification analyses have revealed that "high" fashion involved consumers, are much heavier clothing fashion buyers across most, if not all, classes of men's wear, than their less fashion involved counterparts.

TABLE 4

## Fashion Involvement and Casual Slacks Buying

## 1. Units

Pairs of Casual Slacks <sup>2</sup> purchased in Past Year	Level of Fashion Involvement <sup>1</sup>			Base
	Low	Medium	High	
None	38.8%*	36.3%	14.0%	(369)
1 or 2	41.1	42.8	44.2	(421)
3 or more	20.1	20.9	41.8	(215)
Total	100.0%	100.0%	100.0%	
Base	(541)	(421)	( 43)	

<sup>1</sup>The differences were statistically significant at the .035 level based on X<sup>2</sup> analysis.

<sup>2</sup>Casual slacks were defined as "jeans, corduroy, denim".

\*READ: 38.8 percent of the "low" fashion involved group said they did not purchase any casual slacks during the past year.

## 2. Prices

Average Price Paid per Pair of Casual Slacks <sup>2</sup>	Level of Fashion Involvement <sup>1</sup>			Base
	Low	Medium	High	
Less than \$15.00	51.5%*	38.0%	26.3%	(281)
\$15.00 - \$19.00	34.2	45.1	55.3	(254)
\$20.00 and over	14.2	16.9	18.4	( 99)
Total	100.0%	100.0%	100.0%	
Base	(330)	(266)	( 38)	

<sup>1</sup>The differences were statistically significant at the .006 level based on X<sup>2</sup> analysis.

<sup>2</sup>Casual slacks were defined as "jeans, corduroy, denim".

\*READ: 51.5 percent of the "low" fashion involved group said they paid an average of less than \$15.00 per pair of casual slacks during the past year.

## Conclusions and Implications

The concept of fashion involvement as a summary or cumulative statement of at least five dimensions of fashion adoption-related behavior was defined and researched. An index of fashion involvement was developed and the construct was validated with an independent fashion specific lifestyle/AIO factor analytic technique.

Based on this index, there are major differences among consumers in terms of fashion involvement. The distribution of the male and female populations in the Toronto fashion market reflected a wide range of fashion involvement. Toronto males, however, were lower in fashion involvement than their female counterparts.

The high fashion involved consumer has been the historical interest of fashion researchers, as the drivers and

influentials and legitimitizers of the overall fashion adoption process. This concept has been supported in this paper as a result of a strong and significant relationship identified between fashion involvement and unit and dollar clothing fashion buying behavior.

The findings obtained here strongly suggest that in addition to being the driving force in the fashion adoption process, the highly fashion involved consumers represent an important market as heavy clothing buyers. While, the highly fashion involved group is relatively small vis-a-vis the total population, that group is much larger in terms of proportionate share of clothing fashions purchased. For example, the "high" fashion involved group of suit buying population (4 percent) bought over 10 percent of the suits.

In summary, this paper is part of an on-going research program focusing on fashion adoption. More specifically, this paper has:

1. Developed a fashion involvement index which has face validity and is verified as a construct in independent lifestyle/AIO research.

2. The fashion involvement construct suggests that men and women are different in terms of their fashion involvement.

3. The "high" fashion involved consumer is important to monitor for the fashion industry and particularly the fashion merchant. The "high" fashion involved consumer is the fashion leader in innovativeness, early trial, and interpersonal communication of fashion information. This market segment is also a market target representing a disproportionately high buying segment of the population.

Further research in this program will explore the dynamics of the fashion buying process in greater detail.

## References

- Baumgarten, Steven A. "The Innovative Communicator in the Diffusion Process." *Journal of Marketing Research*, 12 (February 1975), 12-18.
- Grindereng, Margaret P. "Fashion Diffusion." *Journal of Home Economics*, 59 (March 1967), 171-174.
- King, Charles W. "Fashion Adoption: A Rebuttal to the 'Trickle Down' Theory." In Stephen A. Greysier, editor, *Toward Scientific Marketing*, Chicago. American Marketing Association, 1963, 108-125.
- King, Charles W., and L.J. Ring "Retail Fashion Segmentation Research: Development and Implementation." In Bent Stidsen, editor, *Marketing in the 1970's and Beyond*, Edmonton, Canada. Canadian Association of Administrative Sciences, 1975.
- King, Charles W., Tigert, Douglas J. and Ring, L.J., "Contemporary Fashion Theory and Retail Shopping Behavior: A Segmentation Analysis." In the *1975 Combined Proceedings*, American Marketing Association.
- Reynolds, Fred D., and Williams R. Darden. "Why the Midi Failed." *Journal of Advertising Research*, 12 (August 1972), 39-44.
- Sproles, George B., and Charles W. King. "The Consumer Fashion Change Agent: A Theoretical Conceptualization and Empirical Identification." Institute for Research in the Behavioral, Economic, and Management Sciences, Purdue University, Paper No. 433, (November 1973).

References (cont'd)

Summers, John O. "The Identity of Women's Clothing Fashion Opinion Leaders." Journal of Marketing Research, 7 (May 1970), 178-185.

Tigert, Douglas J. "Can a Separate Marketing Strategy for French Canada be Justified: Profiling English-French Markets Through Life Style Analysis." In Canadian Marketing: Problems and Prospects, editor Thompson and Leighton, Toronto, Canada. Wiley Publishers of Canada, Ltd., 1973.

Wells, William D., and Douglas J. Tigert. "Activities, Interests, and Opinions." Journal of Advertising Research, 11 (August 1971), 27-35.

# AN EMPIRICAL EVALUATION OF COMPARATIVE ADVERTISING MESSAGES:

## SUBJECTS' RESPONSES ON PERCEPTUAL DIMENSIONS<sup>1</sup>

R. Dale Wilson (student), The University of Iowa<sup>2</sup>

### Abstract

This paper attempts to evaluate comparative advertising messages which provide little or no factual information to consumers. Subjects' evaluations of simulated comparative advertisements, when compared to other subjects' responses to simulated single-product ads, generally supported the expectation that the comparative messages would be evaluated less favorably than the single-product messages. The results of this study provide preliminary evidence that advertisers should avoid subjective, non-factual comparisons with named competitors.

### Introduction

Of the recent innovations in marketing, perhaps one of the most noticeable is comparative advertising. Contrary to the prohibitive rules of thumb and ethical standards that have existed in the past, some advertisers are now willing to use direct confrontation in their advertising campaigns. By calling their competitor(s) by name and by using side-by-side comparisons, these advertisers believe that they are better able to "prove" the problem-solving utility of their products to existing and potential customers. The popularity of comparative advertising can readily be noted from casual inspection of the print and broadcast media; readers and viewers have recently been exposed to comparative ads for personal care products (e.g., Dial Very Dry antiperspirant vs. Arrid, Sure, and Right Guard), wearing apparel (e.g., No Nonsense pantyhose vs. L'eggs), major appliances (e.g., Magnavox Videomatic TV vs. Zenith and RCA), and automobile warranties (e.g., American Motors vs. Ford, Chrysler, and Chevrolet) among others.

### Background

The door to comparative advertising was formally opened in 1971 when the Federal Trade Commission (FTC) encouraged the major television networks to accept advertisements that included comparisons with (named) competitors' brands. Ulanoff (1975) has indicated that direct comparisons were encouraged by the FTC to provide an acceptable alternative to indirect comparisons using "Brand X" or electronic "beep" sound techniques. Presumably, the FTC's action was motivated by the intuition that comparative advertising would

"...encourage the provision of more factual product information to the consumer and simultaneously to discourage deception by eliminating comparison by innuendo" (Oliver, 1975).

<sup>1</sup>This paper has benefited from the guidance and comments of Professor E. John Kottman (University of Iowa). The author also wishes to thank Gilbert A. Frisbie, Jr. (Bowling Green State University) for his assistance with data collection.

<sup>2</sup>R. Dale Wilson is a Ph.D. candidate in marketing at The University of Iowa.

While the FTC actively supported head-to-head competition using comparative advertising, the criteria for determining fairness, deception and misrepresentation, and substantiation of claims was, for all practical purposes, left for the media themselves to decide. Subsequently, the broadcast media (Christopher, 1974a; 1974b) and the advertising trade associations (Broadcasting, 1974a; American Association of Advertising Agencies, 1974) established a series of self-regulations governing what could, and what could not, be said and done in a comparative ad. These guidelines, coming just before a substantial surge in comparative advertising, were constructed in response to a ruling by the National Advertising Review Board (NARB) that the Schick Flexamatic shaver's advertised claims of superiority over Norelco, Remington, and Sunbeam shavers were in some respects false and were misleading overall (Broadcasting, 1974b).

Despite the good intentions of the media and the self-regulatory boards, the information content of many advertisements using the comparative strategy remains tainted by trivial comparisons that neglect competitive advertising's role as an aid to consumer decision-making. Comparative ads, when they merely compare one brand with another without providing factual information, simply are not useful to the consumer (Chevins, 1974). Evidence that many comparative advertisements consist of claims that are of questionable worth is provided by the persistent warnings of advertising authorities that abuse of this type of advertising must be halted. If action is not taken within the industry itself, the image and credibility of advertising will surely be further damaged, and ultimately the privilege of comparative advertising might be withdrawn by the FTC.

Further evidence of the subjectivity of some comparative ads is the large number of complaints and challenges from those companies who feel that they have been unfairly treated in their competitors' advertisements. Indeed, the volume of protests is so great that one major television network recently issued a set of procedural guidelines specifically designed to expedite the handling of these objections (Advertising Age, 1975). Many of these comparative ads are so obviously subjective (if not false and/or misleading) that their claims defy substantiation.<sup>4</sup>

The managerial and public policy questions stemming from the type of comparative advertising described above are many and varied. Putting aside the obviously important issues regarding the regulation of marketing practices, let us turn to an evaluation of the managerial ramifications of using comparative ads that provide consumers with little or no factual information on which they can base their purchasing decisions. Is this type of advertising not likely to contribute to the negative image of advertising? Will this type of advertising have a negative impact upon consumers' evaluations of products

<sup>3</sup>See comments by Chevins (1974), McMahan (1974), Light (1975), Schneider (1975), and Tyler (1975).

<sup>4</sup>See Cohen (1974, p. 11) for an enlightening discussion of the legal criteria with regard to unsubstantiated claims and the determination of fairness in advertising.

that are advertised in this way? Will advertisements that use subjective and non-factual comparisons lead to a decreased level of credibility and trustworthiness for those companies who advertise their products in this way? Based upon the beliefs of several members of the advertising community<sup>5</sup> it was expected that these questions could be answered in the affirmative. Consequently, a study was designed to explore the differences in consumers' perceptual and attitudinal responses towards comparative and single-product advertising messages.

#### Methodology

In order to assess the impact of subjective claims in comparative advertising, 80 male and female undergraduates enrolled in introductory marketing classes at the University of Iowa were asked to evaluate eight different advertising messages on seven different response dimensions. Each subject was presented with a booklet containing either eight randomly ordered single-product messages or eight randomly ordered comparative messages.

Each simulated advertisement, which consisted of typed headlines and copy, contained a message for a well-known brand in each of the following product classes: automobiles, bath soap, beer, cat food, credit cards, deodorant, mouthwash, and toothpaste.<sup>6</sup> In the comparative ads a competitor's brand name was mentioned while the single-product ad was modified by deleting the competitor's brand name or by replacing the competing brand name with a phrase such as "any other brand" or "other brands." For example, the comparative advertisement for toothpaste was as follows:

ONLY A DENTIST CAN GIVE CHILDREN A BETTER  
FLUORIDE TREATMENT THAN COLGATE.

A kid in the cavity-prone years can wind up with a mouthful of cavities. Eleven is average between five and fifteen. But a Colgate fluoride treatment every day can help prevent some of those cavities. The cavity fighters in Colgate make it more effective than any other toothpaste--even Crest.

Colgate with MFP fluoride helps strengthen teeth against decay. Remember, only your dentist can give teeth a better fluoride treatment.

In the single-product ad, the phrase "even Crest" was merely deleted from the copy. The length of the comparative ads ranged from 49 to 117 words and averaged 74.1 words while the single-product ads averaged 1.4 words longer.

The subjects were told that the advertisements had appeared or were scheduled to appear in the print and/or broadcast media. In reality, only four of the messages (for a brand of automobile, cat food, credit card, and toothpaste) were modified from actual printed advertisements; the remaining ads were constructed specially for the purpose of the experiment. Subjects who received single-product messages were not aware that other subjects were evaluating comparative messages, and vice versa.

<sup>5</sup>See footnote 3.

<sup>6</sup>Product classes and brands were chosen on the basis of a product usage and brand familiarity survey administered to a similar group of undergraduates three months prior to the main study.

Following each simulated advertisement was a page of 7-point scales. Subjects were asked to rate each advertisement on four dimensions (i.e., extremely informative--not at all informative; extremely believable--not at all believable; extremely offensive--not at all offensive; and extremely interesting--not at all interesting), each advertised product on two dimensions (i.e., definitely would change my view towards the product--definitely would not change my view towards the product and extremely high quality--extremely low quality), and each sponsoring company on one dimension (i.e., extremely trustworthy--not at all trustworthy). Thus, seven response measures were collected on each of eight advertisements for every participating subject.

The methodology discussed above, in which each subject evaluates either comparative or single-product advertisements for eight different products forms an 8 x 2 two-factor design with repeated measures on one factor.<sup>7</sup> Briefly, this design is a two-factor (A X B) design in which all subjects are exposed to each of a group of A treatments in combination with any one B treatment, but each group of Ss is exposed to a different B treatment. In our case, one group of Ss was exposed to a series of eight single-product ads, and the other group was subjected to a series of eight comparative ads. This mixture of the A X B factorial design and the treatments X subjects (A X S) design contains treatment comparisons which are mixed--some are inter-subject (between) comparisons while others are intra-subject (within) comparisons. While the two-factor design with repeated measures on one factor is not widely used in consumer research, it offers an important advantage over alternative designs in that a limited number of subjects can be more effectively utilized. Even though carry-over effects are a potential problem for most behavioral applications, this problem can usually be minimized by randomizing the A treatments.

#### Results and Discussion

Each of the seven response variables was considered independently for the purposes of analysis. However, it should be pointed out that inspection of the response variables' correlation matrices shows a relatively high degree of correlation among the seven response variables. But because the correlation coefficients vary widely depending upon the product and type of advertisement being considered, this intercorrelation problem is difficult to handle. Factor analysis of the seven response variables does not provide conclusive evidence as to which, if any, of the response variables should be eliminated from further study. The factor loadings, as well as the number of emerging factors, vary depending upon the product and type of advertisement being analyzed. Consequently, caution must be exercised when interpreting the results of this study.

The results of the analysis of variance stemming from the two-factor design with repeated measures on one factor are presented in Table 1. Each analysis of variance was computed for 75 subjects (35 Ss in the single-product ad treatment and 40 Ss in the comparative advertisement treatment). Five questionnaires were dropped from the analysis because they were improperly completed. The measurement scales were numbered so that 7 indicates a positive response (e.g., extremely informative) and 1 indicates a negative response (e.g., not at all informative).

<sup>7</sup>This type of "mixed" design is known in the experimental design literature under a variety of names. It is discussed in detail by Lindquist (1953, pp. 267-73) and Winer (1971, pp. 518-39).

TABLE 1

## Summary of Analysis of Variance

Source	df	Response Variable							
		Informative Value		Believability		Offensiveness		Interest Produced	
		Mean Sq.	F-Ratio	Mean Sq.	F-Ratio	Mean Sq.	F-Ratio	Mean Sq.	F-Ratio
Between Subjects	74	4.58		4.10		9.83		5.10	
Type of Ad	1	5.79	1.27	13.36	3.37 <sup>1</sup>	89.07	10.18 <sup>2</sup>	8.39	1.66
Error (between)	73	4.56		3.97		8.75		5.05	
Within Subjects	525	2.37		1.67		1.73		1.95	
Products	7	75.44	54.76 <sup>3</sup>	33.70	27.29 <sup>3</sup>	15.61	10.26 <sup>3</sup>	25.16	15.36 <sup>3</sup>
Products/Ads Interaction	7	1.49	1.08	1.12	0.91	3.04	2.00 <sup>1</sup>	1.74	1.06
Error (within)	511	1.38		1.23		1.52		1.64	
Total	599	2.64		1.97		2.73		2.34	

Source	df	Response Variable					
		Change View		Quality		Trustworthiness	
		Mean Sq.	F-Ratio	Mean Sq.	F-Ratio	Mean Sq.	F-Ratio
Between Subjects	74	6.06		2.73		3.95	
Type of Ad	1	2.61	0.43	0.04	0.01	1.14	0.36
Error (between)	73	6.11		2.77		3.98	
Within Subjects	525	2.42		1.21		1.33	
Products	7	27.11	13.07 <sup>3</sup>	14.55	14.16 <sup>3</sup>	18.46	16.89 <sup>3</sup>
Products/Ads Interaction	7	3.21	1.55	0.86	0.83	1.61	1.47
Error (within)	511	2.07		1.03		1.09	
Total	599	2.87		1.39		1.65	

<sup>1</sup>p<.10<sup>2</sup>p<.005<sup>3</sup>p<.001

As can be seen from the table, the sums of squares for the type of simulated advertisements (i.e., single-product vs. comparative ads) is a "between" subjects effect and the sums of squares for products and the products/ads interaction are "within" subjects effects. The relevant F-ratios are formed by dividing the mean squares by the appropriate error term--the type of ad F-ratio is obtained by using the "between" error mean square while the products and products/ads interaction F-ratios are determined by using the "within" error mean square as the denominator.

The results of the ANOVA tables indicate a highly significant ( $p < .001$  in all cases) main effect for the product treatment. Because of the diversity of the products under consideration, this was to be expected. Of primary interest, however, is the main effect of the type of ads used in the study. Table 1 indicates that for response variables believability and offensiveness, the main effect is significant ( $p < .10$  and  $p < .005$  respectively). The significant F-ratios are indicative that the treatment groups differentially responded to the believability and offensiveness dimensions based upon whether they read single-product messages or comparative messages. For the remaining five response means, the main effects of the type of ad were not statistically significant ( $p > .10$ ).

For the believability and offensiveness response dimensions, the differences between the type of ad treatment means were evaluated via Scheffé's (1953) test for multiple comparisons. On the believability dimension, significant differences were found for bath soap ( $p < .05$ ), deodorant ( $p < .10$ ), and mouthwash ( $p < .10$ ). For offensiveness, significant differences between the type of advertisement treatment means for the brands of automobiles

( $p < .05$ ), bath soap ( $p < .01$ ), beer ( $p < .10$ ), cat food ( $p < .10$ ), credit cards ( $p < .01$ ), and deodorant ( $p < .01$ ) were found to exist.

The nature of the differences between response variable means for each product can be more clearly examined in Table 2, which presents differences between means for the single-product and comparative advertisement groups. Of the 56 response differences, 39 (or 69.6%) are in the expected direction. Of these 39 in which the single-product ad is rated more positively than the comparative advertisement, 16 (or 41.0%) are significant at  $p < .10$  or less.

The results of the analysis of variance and the t-tests between means lends support to the notion that the respondents generally evaluated subjective comparative advertising messages more negatively than they did those ads which did not mention a competitor's product by name. This finding must be qualified, however, by stating that the results of this study seem to be product (or advertisement) specific. In addition, the differences in response means across all products on the ability of the ads to change the respondents' views towards the product, the quality of the product, and the trustworthiness of the sponsor appear to be more random than systematic. From Table 2, however, it is clear that for all advertising messages except mouthwash and toothpaste, respondents differentially evaluated the types of ads on the informative value, believability, offensiveness, and/or interest produced dimensions.

#### Conclusion

The results of the study presented in this paper seem to

TABLE 2

Differences Between Mean Responses for Single-Product and Comparative Advertising Messages

Advertising Message	Informative Value	Believability	Offensiveness	Interest Produced	Change View	Quality	Trustworthiness
Automobiles	.504 <sup>1</sup>	.029*	.818 <sup>4</sup>	.311	.364*	.007*	.357*
Bath Soap	.225	.332	1.357 <sup>5</sup>	.468 <sup>1</sup>	.036*	.171	.486 <sup>2</sup>
Beer	.196	.564 <sup>2</sup>	.664 <sup>2</sup>	.275	.225	.268*	.025*
Cat Food	.536 <sup>1</sup>	.229	.675 <sup>2</sup>	.700 <sup>2</sup>	.279	.193*	.296
Credit Cards	.371*	.071	1.075 <sup>4</sup>	.164*	1.007*	.093	.068*
Deodorant	.279	.668 <sup>3</sup>	1.079 <sup>4</sup>	.418 <sup>1</sup>	.121*	.132	.379 <sup>1</sup>
Mouthwash	.125	.436 <sup>1</sup>	.421	.029	.143	.250	.229
Toothpaste	.082	.121	.089	.139*	.175*	.311*	.161*

\*Indicates that the mean difference was not in the expected direction.

<sup>1</sup>  $p < .10$  (one-tailed test)

<sup>2</sup>  $p < .05$

<sup>3</sup>  $p < .01$

<sup>4</sup>  $p < .005$

<sup>5</sup>  $p < .001$

indicate that advertisers may be using comparative ads in a less-than-desirable way. Current advertising campaigns serve as evidence that some advertisers believe that brand comparison, in and of itself, will lead to a more favorable image in the consumer's eye. However, the empirical study discussed earlier suggests that in cases where brand comparisons provide little or no factual information, comparative advertising offers no advantage. In fact, on the dimensions of believability and offensiveness (and to a lesser extent for informative value and interest produced), the two-product comparative ads are evaluated more negatively than single-product advertisements. When one brand does not offer a clear-cut, substantiable advantage over another brand, it seems that consumers should be encouraged to compare products for themselves instead of having non-factual comparisons prepared for them.

The study discussed in this paper adds credence to the suggestion that comparative ads may further damage the institution of advertising. While recent studies have documented the negative image of advertising,<sup>8</sup> comparative advertising may contribute even more to that image. This would seem to be especially true if this type of advertising continues to use comparisons that do not assist consumers in their decision-making processes.

At the present time, an overall evaluation of comparative advertising suffers from insufficient theoretical development<sup>9</sup> and empirical research. Although this paper summarizes one attempt at an empirical evaluation of the content of comparative ads, its findings must be interpreted with caution. The study was conducted on a limited budget, in one geographical area, and used a limited sample size and a small number of advertisements. Additional research must be conducted in order to determine generalized patterns of consumer response to this unique form of advertising.

#### References

Advertising Age, "NBC Sets Guidelines for Comparative Ad Complaints," 46(January 6, 1975),8.

American Association of Advertising Agencies, Policy Statement and Guidelines for Comparative Advertising (New York: AAAA, April, 1974).

Broadcasting, "NBC Puts on Paper its Rules for Comparative Ads," 86(February 4, 1974a),41.

Broadcasting, "Code Review Board Adds More Teeth to TV Rules for Comparison Ads," 87(October 7, 1974b),61.

Business Week, "The Public is Wary of Ads, Too," (January 29, 1972),69.

Anthony Chevins, quoted in "Care in Comparison Urged...", Advertising Age, 45(August 5, 1974),31-2.

<sup>8</sup> See, for example, the results of a study by Warwick and Legler, Inc., summarized in Business Week (1972).

<sup>9</sup> Theoretically, comparative advertising may be considered to be a special case of two-sided advertising (see Karp, 1971). If so, McGuire's (1964) inoculation theory may prove useful as a basis for a theoretical evaluation of comparative advertising. In addition, Wilkie and Farris (1974) have provided a theoretical framework for comparative advertising based on Lavidge and Steiner's (1961) "hierarchy of effects" model.

Maurine Christopher, "NBC Spells Out New Formal Guides for Comparative Spots," Advertising Age, 45(January 28, 1974a),1+.

Maurine Christopher, "ABC Comparative Rules Stress Test Procedures," Advertising Age, 45(March 18, 1974b),1+.

Dorothy Cohen, "The Concept of Unfairness as It Relates to Advertising Legislation," Journal of Marketing, 38(July, 1974),8-13.

Robert E. Karp, "Creative Advertising Strategies: The Two-Sided Approach," Business and Society, 12(February, 1971),18-25.

Robert J. Lavidge and Gary A. Steiner, "A Model for Predictive Measurements of Advertising Effectiveness," Journal of Marketing, 25(October, 1961),59-62.

Lawrence Light, quoted in "Underdog Advertiser Wins in 'Naming Names': BBDO," Advertising Age, 46(March 10, 1975),56.

E. F. Lindquist, Design and Analysis of Experiments in Psychology and Education (Boston: Houghton Mifflin, 1953).

William J. McGuire, "Inducing Resistance to Persuasion: Some Contemporary Approaches," in Leonard Berkowitz (ed.), Advances in Experimental Social Psychology (New York: Academic Press, 1964),191-229.

Harry Wayne McMahan, quoted in "Advertisers' Name-calling Could Hurt, McMahan Says," Advertising Age, 45(February 18, 1974),72.

Robert E. Oliver, quoted in "Open Way for Comparative Ads, Canada Meet Told," Advertising Age, 46(May 12, 1975),6.

Henry Scheffé, "A Method for Judging All Contrasts in the Analysis of Variance," Biometrika, 40(June, 1953), 87-104.

Alfred Schneider, quoted in "ABC Censor Raps Trend to Naming Rivals in Ads," Advertising Age, 46(March 31, 1975),98.

William D. Tyler, "Comparative Advertising: A Powerful Selling Tool When It is Not Abused," Advertising Age, 46(April 21, 1975),58+.

Stanley M. Ulanoff, "Comparison Advertising: An Historical Retrospective," working paper (Cambridge, Mass.: Marketing Science Institute, February, 1975).

William L. Wilkie and Paul Farris, "Comparison Advertising: Issues and Prospects," working paper (Cambridge, Mass.: Marketing Science Institute, August, 1974).

B. J. Winer, Statistical Principles in Experimental Design, Second Ed. (New York: McGraw-Hill, 1971).



## RESEARCH METHODOLOGY AND ADVERTISING SUBSTANTIATION

Robert A. Mittelstaedt, University of Nebraska  
Nils-Erik Aaby (student) University of Nebraska

### Abstract

As more consumer research is used to substantiate advertising claims in FTC proceedings, methodological questions become important. Full utilization of research results depends on correctly interpreting the errors (Type I and Type II) inherent in accepting or rejecting hypotheses and some estimate of the relative costs associated with those errors.

### Problem

The use of research results to substantiate advertising claims which have been challenged as deceptive is not new. However, the Federal Trade Commission has been formalizing this process through its advertising substantiation program and the relationship between research and product claims has taken on several new aspects.

First, many more claims are being scrutinized and the frequency of calls for substantiation has increased. Second, under the new "streamlined and tightened" rules of January, 1974, the response time has been shortened from 60 to 30 days (FTC News Summary, 1974a) with a proposal to exclude from adjudicative proceedings any material arriving after the deadline (FTC News Summary, 1974b). No doubt this reflects recent Commission decisions that research supporting a claim must be performed prior to the making of a claim (Pfizer, 1972).

Third, the type of research that would be acceptable to support a given claim appears to have been broadened. In 1972 the Commission defined a "scientific test" as follows:

In our view a scientific test is one in which persons with skill and expertise in the field conduct the test and evaluate its results in a disinterested manner using testing procedures generally accepted in the profession which best ensure accurate results. This is not to say that respondent always must conduct laboratory tests. The appropriate test depends on the nature of the claim made. Thus a road or user test may be an adequate scientific test to substantiate one performance claim, whereas a laboratory test may be the proper test to substantiate another claim. Respondent's obligation is to assure that any claim it makes is adequately substantiated by the results of whatever constitutes a scientific test in those circumstances (Firestone, 1972, 463).

More recently, the Commission appears to have put the "hearsay issue," which has long been a problem when using consumer research as legal evidence, to rest:

Turning to the test marketing reports in this record, we must dismiss any contention that the FTC is bound to reject these consumer surveys as inadmissible hearsay. The Commission has on numerous occasions considered the question of the admissibility of surveys which are obviously hearsay, and it is well settled that such surveys will be admitted for the truth of the matters asserted when

it is demonstrated that they are reasonably reliable and probative (Bristol-Myers, 1975).

Further, the Commission has used "academic" research as an input into a decision regarding a proposed ban on premium ads on children's television programs (Advertising Age, 1975).

Finally, while the Commission has apparently broadened its definition of scientific evidence, it has given considerably more attention to methodological details. (Alberto-Culver, 1973; Gillette, 1973) For example, the Commission ordered the Gillette Company to substantiate a claim that a deodorant had "A light, clean scent. Not perfumey or chemical. A scent that comes from real, natural ingredients. Not a lot of artificial ones" (Gillette, 1973) by documentation that was to include, but not be limited to (Gillette, 1973):

- A. A definition of such terms as "light," "clean," etc.
- B. Survey or experimental evidence that a substantial majority of individuals perceived Right Guard as light, clean, etc. to include:
  1. complete explanation of methodology.
  2. description of subjects, method of recruitment, size and generalizability of sample.
  3. experience and training of experimenters.
  4. list and descriptions of products compared.
  5. survey respondents' knowledge of purpose.
  6. statistical designs and techniques of statistical inference including assumptions necessary for tests of significance and specification of alpha level.
  7. data necessary for replication.
  8. copies of all questionnaires.
  9. indication of reliability and validity of all instruments used.
- C. List of ingredients of product, by percentage, and their function.

All of this suggests that in cases involving the substantiation of advertising claims, research of the sort familiar to consumer researchers will play an increasingly important role. It further suggests that the Commission, as it seeks to specify certain methodological matters, will be less able to reject research evidence as "not germane" and the resulting evidentiary issues will increasingly turn on methodological questions. The purpose of this paper is to discuss some of the logical, methodological and legal issues involved in using research as evidence in cases involving the substantiation of advertising claims.

### A Hypothetical Example

Assume Firm "X" wishes to claim that its product "A" is better than "B." In this example, "B" may be a previous product, a competitive product or some established standard. Regardless of the referent, the claim is

essentially comparative and presumably verified by showing "A" to be "better than 'B'" along some measurable dimension(s). If Firm "X" makes such a claim in its advertising, the present interpretation of the substantiation program is that "X" must have some basis for the claim, presumably research results. If called upon to substantiate, the respondent ("X") submits its research and the Commission decides whether or not it supports the claim.

At this point, there are at least two problems, one conceptual and the other legal, which will be noted but not discussed. First, one may question whether the truth of a specific claim is really the central issue in determining "deception," an issue discussed elsewhere (Developments, 1967; Gardner, 1975). Second, the whole procedure may be viewed as shifting the burden of proof to the advertiser, which some authorities view as the "fatal flaw" in the entire ad substantiation program (Gellerhorn 1969, 1971; Rosden, 1974). However, the Commission is requiring research results from advertisers using it to determine the existence of deception.

Whether an advertisement is analyzed from the standpoint of unfairness or deception, however, the standard for evaluating the substantiating material and the test which is applied is the same -- does the substantiation provide a reasonable basis to support the claim. Essentially, this is a factual issue to be formulated in the context of circumstances present in each case (National Dynamics, 1973, 549).

Given this, two general areas, measurement and statistical inference, are discussed before attention is turned to the matter of research evidence in the larger context of the Commission's decision making process.

#### Measurement Issues

No doubt the dimensions along which "A" is supposedly better than "B" will determine the type of measurement used to support the claim. In a general sense three situations seems applicable. First, the measurement may be "physical" or "mechanical," e.g., "A" tires stop faster than "B" or it takes greater pressure to break "A" than "B." Second, the measurements may consist of the observations of some "expert" (the typical clinical evidence), e.g., "A" produces fewer cavities or relieves some set of symptoms better than "B." Third, the perceptions, opinions or attitudes of consumers may be measured, e.g., "A" smells fresher or tastes better than "B."

While all of these measurements lead to some "numbers" and, presumably, to a "test statistic," there are a few differences among these three situations worth noting. In general, the problem of finding a measure which both Firm "X" and the FTC agree is pertinent to the claim is probably most acute in the third case (consumer research). The consumer research case is also most likely to involve any possible issues of "hearsay evidence" that may remain, in addition to the problem of concealing the identity of respondents and the potentially conflicting goals of "full disclosure" and "professional ethics." Furthermore, the nature of the data generated in consumer, and possibly clinical, research will be more likely to lend themselves to analysis by non-parametric methods which are generally less powerful. (The issue of "power" will be discussed later.) Finally, while not exactly a measurement issue, each of the cases described above has its own peculiar sampling problems of defining the relevant "universe" and actually achieving a random selection process.

#### Statistical Inference

Formed as a research question, the support for the claim that "'A' is better than 'B'" would most likely rest on a test capable of rejecting the null hypothesis that "'A' is not better than (or equal to) 'B'." in favor of the alternative. From the data, a test statistic is computed (t, F, Chi-square, etc.) and, if its magnitude indicates that the test results are sufficiently unlikely under the assumption of the null hypothesis, the null is rejected and the result is said to be significant. Of course, unlikely things do occur and there is a chance of wrongly rejecting the null hypothesis (Type I error) or, if non-significant, a chance of wrongly accepting the null (Type II error.) Type I error is "controllable" by the researcher through the selection of the alpha level and "good practice" usually dictates that it is set in advance of the test and at a fairly low, conventional level. Controlling Type II error (beta) is also possible, but more difficult, since it can be computed only against a specified and definite alternative.

If the Commission accepts the results of the test as establishing substantiation when the null hypothesis is rejected (the results are significant) and failing to establish substantiation when the null is not rejected (the results are not significant), they assume the "risk burden" of each type of error in their decision. Table 1 shows this schematically. If the null hypothesis is rejected and the Commission concludes the claim is valid, the Commission runs the risk ( $p=\alpha$ ) of allowing a false claim to stay on the market. On the other hand, if the test is non-significant and the Commission concludes that the ad is unsubstantiated and must be removed, they risk ( $p=\beta$ ) taking a true claim off the market.

TABLE 1

Research Decision	Commission Decision	True State in Population	
		(A no better than B)	(A better than B)
Reject $H_0$	claim true	Type I Error	correct decision
Accept $H_0$	claim false	correct decision	Type II Error

Therefore, using research findings as major inputs to the Commission's decision making in cases involving the substantiation of advertising claims raises at least two issues: the interpretation of statistical significance and the relative costs of Type I and Type II errors.

#### Statistical Significance and "Facts"

As anyone who has ever taught or taken a statistics course knows, the phrase "statistical significance" has a rather special meaning. As Hays puts it:

. . . all that a significant result implies is that one has observed something relatively unlikely given the hypothetical situation, but relatively more likely given some alternative situation. . . . Statistical significance is a statement about the likelihood of the observed result, nothing else. It does not guarantee that something important, or even meaningful, has happened (Hays, 1973, 384).

There appear to be few cases in which the Commission or the Courts have confronted the necessity of interpreting a "statistically significant result" or "non-significance," although some substantiation data appear to lack significance (Aaby, 1975). Such an interpretation should recognize both the probabilistic nature of any conclusion drawn and the matter of "effect size."

Looking first at the probabilistic nature of any conclusion drawn from a statistical test, the Commission and the Courts seem to have taken the results of particular tests as "fact." In *Pfizer* (1972), an expert witness attested to the statistical significance of the results of an efficacy test of a sunburn remedy and the administrative law judge accepted the claim that the product "worked." The Commission skirted the issue of interpretation, but before rejecting the finding as evidence (because the performance of the test had not preceded the making of the claim) noted:

The nature and intricacy of the debate on the adequacy of this test leads to the view that the Commission's role should simply be one of attempting to determine the existence and general quality of the tests and a threshold determination as to the reasonableness of reliance thereon, rather than an attempt to conclusively determine the adequacy of the tests (Pfizer, 1972, 67).

In two earlier cases the Commission and, eventually, the Courts were faced with results which were "in the right direction" but failed to achieve statistical significance. In *FTC vs. Country Tweeds* (1964), the failure of the respondent firm to mention that test results reproduced for promotional purposes had been "insignificant" was held to be misleading by the FTC and specifically upheld by the court. In *FTC vs. Sterling Drug* (1963), the issue was also what the advertiser had said about a non-significant finding. At issue were the data reproduced in Table 2.

TABLE 2

Pain-Relief Scores in Postpartum Patients  
(Excerpted from DeKornfeld, et. al., 1962)

Time Interval (min.)	Standard Error of Mean	Mean Pain-Relief Scores For Indicated Drugs*					
		P1	Bu	An	StJ	Ex	Ba
15	.124	.26	.60	.72	.76	.90	.96

\* Any 2 means not underscored by the same line are significantly different (p .05). Any 2 means underscored by the same line are not significantly different.

P1 = Placebo                      An = Anacin                      Ex = Excedrin  
Bu = Bufferin                      StJ = St. Joseph                      Ba = Bayer  
   Aspirin                                      Aspirin

The ad in question had conceded that there were no significant differences among the products but went on to say "Nevertheless it is interesting to note that within just 15 minutes Bayer Aspirin had a somewhat higher pain relief score than any of the other products." Although the FTC objected to this statement, the Court found the statement "literally true" and went on to observe:

The fact that the margin of accuracy of the scoring system was .124 -- meaning that the second place drug might fare as well or better over the long run of statistical tests -- does not detract from the fact that on this particular test,

Bayer apparently fared better than any other product in relieving pain within fifteen minutes after its administration (FTC vs. Sterling Drug, 1962, 677).

In short, the Commission accepted the lack of significance as establishing "no difference" as fact while the Court seemed to recognize the probabilistic nature of the conclusion but focused on the specific instance reported in the data. However, the data in Table 1 raise two further questions, both related to effect size and the power of the statistical test employed.

The ability of a test to reject a null hypothesis when it is, in fact, false (i.e., the test's power) is a function of sample size, effect size (the real difference) and the level of alpha (Cohen, 1969). Assuming that there is some effect (Bayer really is modestly more effective in fifteen minutes) and the level of alpha fixed at .05 (by convention or decree) there is, presumably, some sample size sufficiently large to produce a "statistically significant result." If the Commission's case rests upon the failure to reject a null hypothesis at some given level of alpha, while the respondent's case rests on rejection, the question of sample size becomes a point of contention between the two. Unless some attention is given to the question of how much "A" must be better than "B," in an adversary proceeding the Commission would maximize the number of its enforcement orders by insisting on small samples while the respondents would always prefer to defend with large samples. When the advertiser has made a specific claim ("A is 25% better than B") the alternative hypothesis is clear. However, in the presence of an indefinite alternative ("A is better than B"), the search for "truth" may devolve into a battle over semantics or among statisticians (Advertising Age, 1974). The FTC and the Courts have recognized that differences may not be practically significant (Lorillard, 1950) but do not appear to have related this to the issue of a test's power.

Also involved is the relationship between the alpha level and the power of a test. Other factors (effect size and sample size) equal, a test's power diminishes as the size of alpha decreases. Put another way, as the probability of making a Type I error is decreased, the probability of making a Type II error increases. If there are costs associated with each error type, the problem goes beyond the statistical issue to the decision making framework of the Commission.

#### Costs of Type I and Type II Errors

In the final analysis, the consumer's interest is not in the confirmation or disconfirmation of a statistical hypothesis but in the "truth" of a given claim. Ultimately, the statistical "result" is a probabilistic statement and, as such, is evidence but not proof. However, by relying on sample evidence from a properly conducted study, the risks of erroneously allowing a false claim to stand (Type I error) and erroneously removing a true claim (Type II error) are estimable. It also seems plausible that the disutility to consumers associated with either of these errors will vary from situation to situation.

Kaplan (1968) has provided a framework for evaluating the degree of certainty which evidence must possess to make a decision, taking the costs of "wrong decisions" into account. Expressed as an inequality:

$$P > \frac{1}{1 + \frac{D_g}{D_i}} \quad (1)$$

Where: P = the degree of certainty needed to convict or find for the plaintiff.

$D_g$  = the disutility of allowing the guilty to go free (i.e., letting a false ad claim stand).

$D_i$  = the disutility of convicting the innocent, (i.e., removing a true ad claim from the market).

Kaplan observes that in most civil actions ("X" sues "Y") the disutilities of erroneously finding for the plaintiff or the defendant are most likely equal ( $D_g/D_i = 1.0$ ). If so, P becomes .5 and any "preponderance of evidence" for the plaintiff (i.e., there is more reason to believe the plaintiff is right) should be sufficient to find for him. However, in criminal cases, because we hold to the general belief that it is better for the guilty to go free than for the innocent to be convicted ( $D_g/D_i < 1.0$ ) the weight of evidence needed to convict (expressed as the probability of belief, P) should be much closer to 1.0. This is especially true when the penalty is severe (e.g., capital punishment).

In Pfizer (1972), the FTC appeared to believe that the disutility of relying on a false claim ( $D_g$ ) fell exclusively on the consumer while the disutility associated with Type II error ( $D_i$ ) fell only on the producer and, therefore  $D_g$  was always larger than  $D_i$ . While not completely rejecting that notion (at least with respect to  $D_i$ ), the recently announced dismissal of the order against "Dry-Ban" seems to indicate that the Commission recognizes that  $D_g$  will vary from case to case (Bristol-Meyers, 1975). Specifically, that decision implies that greater  $D_g$  is experienced when the audience is particularly vulnerable, when health and safety claims are involved and when serious economic loss results from reliance on a false claim. Further, the decision recognizes that injury to competition is a factor involved in estimating the magnitude of  $D_g$ .

The losses associated with erroneously removing a true claim from the market ( $D_i$ ) are less easily recognized but no less real. If product "A" is truly better than "B," but the producer of "A" is forced to stop claiming its superiority, the losses to the firm are obvious -- lost sales and the costs of presenting a defence. However, at least two other disutilities would result. First, consumers would be denied access to potentially useful information (assuming the claim is not only valid but relevant). It should be noted that in a recent case (Crown Central Petroleum, 1974) the Commission explicitly chose not to consider whether a 10% reduction in auto emissions had "social value" and decided the case on the narrower grounds that the gasoline in question didn't produce "clean air" as claimed. A second disutility is involved if the Commission uses some of its limited "regulatory resources" to achieve a faulty goal, resulting in an opportunity loss associated with the consequent failure to pursue some other deception (Mittelstaedt, 1972). But, however the magnitude of  $D_i$  is estimated and whatever it includes, the imposition of the "ultimate penalty," corrective advertising, always increases  $D_i$  relative to  $D_g$ .

The point is that there are disutilities involved in both types of error. The public is less well off when a false ad is allowed to stand or when a true ad is removed. While these disutilities are difficult to quantify, only the ratio of  $D_g/D_i$  is really important and subjective estimates will probably have to suffice.

## Conclusions

In the decision process, both the disutilities and probabilities of each type of error are important and this brings us back to the question of rejecting null hypotheses. Two possible situations seem worthy of comment.

First, assume that a firm has been claiming that its product "A" is better than "B." As substantiation, it has a research study which it and the FTC agree is a "fair test" and which rejects the null hypothesis ("A" is no better than "B") at the .05 level. On the basis of this test alone the Commission cannot know if the claim is really true but, if they accept this finding as fact, they are taking a known chance (.05) of allowing a false claim to stand. Kaplan's model suggests that the Commission should rely on the results of the test unless  $D_g$  exceeds  $D_i$  by a very wide margin (19 to 1). In cases where the costs of allowing the false claim to stand are relatively large (e.g., safety claims), a smaller level of alpha (e.g., .01) may be appropriate.

Second, assume the test's results fail to reject the null hypothesis that "A" is no better than "B." Here the proper action is less clear for the Commission's "case" rests on the acceptance of the null hypothesis as "fact." The researcher faced with non-significant results is usually told by the statistician to follow the Fisherian advice and "suspend judgment." This advice is sound when the alternative hypothesis is indefinite and the costs of error of either type are unknown or small (Hays, 1973, 355-357). While researchers may be able to suspend judgment, the Commission probably can't (or won't). But, if the power of the test used in the research can be calculated, which rests on the ability to specify some meaningful alternative hypothesis, the risks of Type II error can be compared with the relative disutilities of  $D_g$  and  $D_i$ . As the ratio ( $D_g/D_i$ ) grows smaller, the question of the test's power becomes increasingly important. Since "corrective advertising" increases  $D_i$  relative to  $D_g$ , the power of the test should be a critical determination before this remedy is applied.

Increasingly, studies of the sort familiar to consumer researchers will be used to substantiate advertising claims. As researchers (and consumers) we have a stake in seeing that something positive results from the effort. Questions of measurement validity and the generalizability of samples will, and should, continue to engage our attention but, in the end, research results must be interpreted as evidence for or against specific claims. Full utilization of any study's information in the regulatory process requires an understanding of the probabilistic nature of research results. To focus on Type I error alone will not insure correct interpretation. What is needed is an understanding of both Type I and Type II errors and an appreciation of the relative costs of each.

## References

Nils-Erik Aaby, "American Comparative Advertising: The Problem of Substantiation," paper presented at Annual Meeting of Southwestern Marketing Association, Houston (March 8, 1975).

Advertising Age, 16 December 1974, 14.

Advertising Age, 12 May 1975, 32.

Alberto-Culver Co., et. al., 3 Trade Reg. Rep., ¶ -20,357 (1973).

Bristol-Myers Co., et. al., 3 Trade Reg. Rep.,  
¶ -20,900 (1975).

Jacob Cohen, Statistical Power Analysis for the Behavioral Sciences (New York: Academic Press, 1969).

Crown Central Petroleum Corp., 3 Trade Reg. Rep.,  
¶ -20,790 (1974).

Thomas J. Dekornfeld, Louis Lasagna and Todd M. Frazier,  
"A Comparative Study of Five Proprietary Analgesis Compounds," Journal of American Medical Association,  
182(December 29, 1962), 1315-1318.

Developments in the Law, "Deceptive Advertising,"  
Harvard Law Review, 80 (March, 1967), 925-1164.

Firestone Tire and Rubber Co., 81 FTC 463 (1972).

FTC News Summary (no. 2), 16-31 January, 1974(a), 1.

FTC News Summary (no. 10), 16-31 May 1974(b), 1.

FTC News Summary (no. 23), 30 May 1975, 1.

FTC vs. Country Tweeds, 326 F 2d. 144 (1964).

FTC vs. P. Lorillard, 186 F 2d. 52 (1950).

FTC vs. Sterling Drug, et. al., 317 F 2d. 669 (1963).

David M. Gardner, "Deception in Advertising: A Conceptual Approach," Journal of Marketing, 39(January, 1975), 40-46.

Ernest Gellerhorn, "Proof of Consumer Deception Before Federal Trade Commission," University of Kansas Law Review, 17(1969), 559-585.

Ernest Gellerhorn, "Rules of Evidence and Official Notice in Formal Administrative Hearings," Duke Law Journal, 1(1971), 1-50.

Gillette Co., et. al., 3 Trade Reg. Rep., ¶ 20,340 (1973).

William L. Hays, Statistics for the Social Sciences, (New York: Holt, Rinehart and Winston, 1973).

John Kaplan, "Decision Theory and the Factfinding Process," Stanford Law Review, 20 (June, 1968), 1065-1092.

Robert A. Mittelstaedt, "Consumer Protection and the Value of Information," in M. Venkatesan (ed.), Proceedings of the Third Annual Conference of the Association for Consumer Research, (1972), 101-106.

National Dynamics Corporation, et. al., 82 FTC 488 (1973).

Pfizer, Inc., 81 FTC 23 (1972).

Richard A. Posner, Regulation of Advertising by the FTC, (Washington, D. C.: American Enterprise Institute for Public Policy Research, 1973).

George Eric Rosden, The Law of Advertising (vol. 2), (New York: Mathew Bender, 1974).

## CONSUMER REACTIONS TO COMPARATIVE ADVERTISING

Linda L. Golden, University of Texas at Austin

### Abstract

This study investigates the relative influence of comparative and non-comparative advertising upon purchase intentions. The advertiser's competitive position, claim substantiation and theme are considered, while brand loyalty is controlled. The results of the factorial experiment, employing covariance analysis, indicate that comparative advertisements appear to be no more effective than non-comparative advertisements, except when theme is considered.

### Introduction

Comparative advertising is a recently popularized technique which has generated considerable attention and controversy. The purpose of this paper is to briefly describe the background and controversies surrounding comparative advertising and to discuss an empirical investigation of the impact of comparative advertising upon the consumer. For purposes of this investigation, comparative advertising may be defined as any form of paid promotion which:

1. compares the advertised brand to one or more specifically named or recognizably referred to brands within its generic product class;
2. with respect to one or more specific and/or over-all product characteristics.

The first criteria specifies that brands must either be specifically named or referred to in a manner which would make the brands recognizable to the consumer. This criteria excludes any advertising which does not make clear to the audience the brands against which the comparison is made. The second criteria in the definition allows inclusion of advertisements which either compare specific product characteristics or state that the advertised brand is "better" over-all than certain other identifiable brands.

### Background and Controversy Surrounding Comparative Advertising

Until recently the primary form of brand comparisons used in the consumer advertising field were Brand X euphemisms. Although comparisons with named competitors have been a long-established personal selling technique for consumer products, it was not until 1968 when American Motors compared the Javelin to the Mustang that direct brand comparisons began frequenting the print media. The use of named competitors has been slow to evolve in electronic media presentations of consumer products. However, due to the relaxation of traditional advertising taboos, the frequency of direct brand comparisons in consumer advertising has increased dramatically in the last two years.

This increase may be largely attributed to endorsements of comparative advertisements by the Federal Trade Commission and special interest groups. In 1972, the FTC issued a statement calling for advertisers to name competing brands in their advertisements as an alternative to the Brand X euphemism (Bradshaw, 1974). In conjunction with their request to advertisers, the FTC asked

major television networks to change their policies prohibiting explicit mention of competitors in advertisements and to accept these otherwise suitable commercials on a trial basis as a means of informing the consumer about who competes against whom and how (Christopher, 1974).

Other groups, including consumerism advocates, have also cited benefits of comparative advertising in their endorsements. Joan Bernstein, Assistant Director of the Bureau of Consumer Protection, has stressed the potential of comparative advertising for delivering information not previously available to consumers (Grant, 1973). And although the American Association of Advertising Agencies has previously publicly discouraged the use of comparative advertising, a recent policy statement recognizes that comparative advertising provides the consumer with needed and useful information when used truthfully and fairly (Tannenbaum, 1974).

In spite of the benefits acclaimed to comparative advertising and the endorsements, some groups have begun to evaluate comparative advertising for negative side-effects. Jack Roberts, of Ogilvy and Mather, expressed concern that with comparative advertising the environment could become one of claims, counterclaims and contradictions resulting in increased opportunities for misleading or deceptive claims (Roberts, 1973). Both the Television Code and the Radio Code imply recognition of the potential of comparative advertising to be deceptive and misleading. The Television Code states that, "Advertising should offer a product or service on its positive merits and refrain by identification or other means from discrediting, disparaging or unfairly attacking competitors, competing products, other industries, professions or institutions" (Szybillo and Heslin, 1973). And the Radio Code states that "any identification or comparison of a competitive product or service, by name, or other means, should be confined to specific facts rather than generalized statements or conclusions unless such statements or conclusions are not derogatory in nature" (Ulanoff, 1972).

The American Broadcasting Company and the National Broadcasting Company have now developed formal guidelines for comparative commercials to help avoid the possibility of deceptive and misleading information. Although the guidelines of both networks are very similar, ABC places stress on the test procedures for comparative claims. Thus, the controversy over the deceptive nature of comparative advertising has led to more concern over an already scrutinized advertising technique, claim substantiations.

With regard to comparative advertising, several things seem apparent. First, public policy makers, media representatives, and other special interest groups have taken an increased interest in the use and impact of comparative advertising. The result is that many groups have endorsed comparative advertising on the assumption that it will provide the consumer with more and better information than will traditional advertising. Other groups, however, have expressed concern over the ability of comparative advertising to convey deceptive and misleading information to the consumer. Second, consumers have experienced and will probably continue to experience an increase in the use of comparative advertising

strategies. And finally, since the use of comparative advertising has only recently been accepted, advertisers and public policy makers have virtually no information concerning its impact on the consumer. The endorsements and criticisms of comparative advertisements are a result of hypotheses which have neither been tested nor validated, and no one has asked if comparative advertising is effective for the advertiser.

As advertisers are increasingly taking advantage of the fact that they can now employ comparative advertising, it is important that the relative impact of comparative and noncomparative advertising upon the consumer be empirically investigated. Since the ultimate objective of a firm's advertising is to elicit a purchase response from the consumer, the general purpose of this research is to investigate the relative effects of comparative and non-comparative advertising upon purchase intentions. Further, this investigation studies the effects of the communications with respect to other variables which may affect consumer reactions to comparative advertisements. Specifically, the effects of the consumer's degree of brand loyalty toward the advertised brand, degree of brand loyalty toward the competing brand in a comparative advertisement, the competitive position of the advertised brand and claim substantiation are considered in this investigation. The influence of three different but parallel copy themes is also investigated in order to eliminate copy testing of one theme.

#### Methodology

A four part questionnaire was administered to 594 university juniors and seniors. The first section of the questionnaire obtained information concerning the respondent's degree of brand loyalty toward alternative brands of deodorants. The respondent was asked to report what percentage of his deodorant purchases during the last year were Arrid, Right Guard and Sure, respectively. The next question in this section asked the respondent to name any other brands of deodorant he purchased during the last year, and to indicate the percentage of time each was purchased. All brand loyalty questions elicited unaided percentage responses with the stipulation that the percentage must sum to one-hundred percent. The focus of this section was to provide a surrogate measure of brand loyalty toward the advertised brand in the copy manipulation and toward the competing brand (Arrid) in comparative copy treatments.

In the second section of the questionnaire, the respondent was exposed to a fictitious piece of deodorant copy for either Right Guard, Sure, or a fictitious new brand (Secure). The copy explicitly stated the market position of the advertised brand, with Right Guard having the number one market position, Sure occupying the number three market position, and Secure being identified as a new brand. In the comparative copy manipulation, Arrid was always the competing brand and identified as occupying the number two market position. The comparative copy stated that the advertised brand was better than Arrid for exactly two product characteristics: ability to stop wetness and odor. Other product characteristics mentioned in all copy manipulations without a comparison were: goes on dry, keeps you dry, long-lasting, and never stings or stains. One-half of the claims in the copy manipulations were substantiated with the statement, "An independent testing agency's results have proven that..." while the other half of the copy manipulations did not include any claim substantiation statement. Three separate but parallel themes were included in the design to avoid copy testing and to increase the external validity of the study. Thus, the respondent could receive any one of thirty-six different

copy manipulations included in the factorial design (2x3x2x3). There were two levels of advertisement copy type (comparative or non-comparative), three levels of competitive position (new brand, number one brand or number three brand), two levels of claim substantiation (substantiated or unsubstantiated), and three levels of copy theme. Respondents were randomly assigned to treatments and except for the inclusion of the experimental manipulations, all treatments were parallel.

The third section of the questionnaire asked the respondent to indicate how likely he would be to purchase the advertised brand. Purchase intention ratings were measured on a horizontal nine-point scale with the highest extreme labeled "very likely" and the lowest extreme labeled "very unlikely."

The last section of the questionnaire obtained information concerning the relative importance of product characteristics for which brand comparisons were made and other characteristics which were mentioned or not mentioned in the advertisements. Nine characteristics were presented to the respondent: price, scent, stops odor, stops wetness, applies dry, does not stain, does not sting, long-lasting and over-all effectiveness, respectively. Respondents were asked to indicate on a horizontal nine point scale (highest extreme labeled "very important") how important each characteristic was to them when purchasing a deodorant.

The data from the first three sections of the questionnaire were submitted to analysis of covariance (Barr and Goodnight, 1972). Degree of brand loyalty toward the advertised brand and degree of brand loyalty toward the competing brand in the comparative advertisement were the covariates in this analysis. Thus, the influence of brand loyalty upon purchase intention ratings could be controlled. The product characteristic importance ratings were submitted to analysis of variance (Dixon, 1970).

#### Validation Study

One limit to the external validity of this study is the use of copy manipulations. Copy manipulations were used to limit the amount of stimuli the respondent was exposed to and to allow as clean an attribution as possible that the results were a reaction to manipulations rather than to other features of a full advertisement. However, it is difficult to assume that consumers would react to full comparative advertisements in the same way as they would respond to comparative copy manipulations. For this reason, a validation study was conducted with full advertisements currently in print media to investigate differences in consumer reactions to comparative and non-comparative advertisements and copy.

The design of the questionnaire was similar to that used for the major study. The questionnaire obtained exactly the same information concerning the respondent's degree of brand loyalty toward both the advertised brand and the competing brand, as well as purchase intentions toward the advertised brand. The experimental manipulation consisted of one of two forms of an advertisement appearing in a local Sunday newspaper. The comparative advertisement for Dial deodorant was constructed so that the comparisons appeared in the upper left corner and could easily be eliminated for the non-comparative treatment. The validation study was administered to 77 university juniors and seniors with respondents randomly assigned to either the comparative or non-comparative treatment. The data were submitted to analysis of covariance (Barr and Goodnight, 1972).

## Results

A summary of the analysis of covariance for purchase intention ratings toward the advertised brand appears in Table 1. Although all interactions were included in the analysis, only the results of the significant interactions and main effects are included in the summary table. The terms included in the model explain thirty-eight percent of the variation of purchase intention ratings. The probability of explaining this amount of the variation by chance is less than one percent.

TABLE 1

Analysis of Covariance for Purchase Intentions

Source	d.f.	Sums of Squares	F-Ratio
Regression	37	1522.18	9.05 <sup>a</sup>
Loyal <sup>d</sup>	1	1006.22	221.47 <sup>a</sup>
Cloyal <sup>e</sup>	1	11.65	2.56
Ad	1	9.41	2.07
Comp	2	21.43	2.36
Claim	1	1.64	.36
Theme	2	4.02	.44 <sup>b</sup>
Ad X Theme	2	37.80	4.16 <sup>b</sup>
Comp X Theme	4	82.95	4.56 <sup>a</sup>
Error	556	2526.15	

<sup>a</sup>  $p \leq .01$

<sup>b</sup>  $p \leq .01$

<sup>c</sup>  $R^2 = .38$ ; Intercept = 2.82

<sup>d</sup> Degree of brand loyalty toward the advertised brand

<sup>e</sup> Degree of brand loyalty toward the competing brand

There is no significant main effect of advertisement copy type.<sup>1</sup> Thus, the purchase intention ratings elicited by the comparative advertisement are not significantly different from those of the non-comparative advertisement. There is also no significant main effect of competitive position,<sup>2</sup> claim substantiation or theme.

<sup>1</sup> The correlation coefficient for degree of brand loyalty toward the competing brand and advertisement copy type is .35. The effect of this may be downward biased F-values for advertisement copy type, thus these results should be interpreted with caution. For the comparative treatment, the partial correlation of the degree of brand loyalty toward the competing brand with purchase intentions is .08. Since there was no competing brand in the non-comparative copy, all respondents in this treatment received a zero for degree of brand loyalty toward the competing brand.

<sup>2</sup> Respondents receiving an advertisement for the fictitious brand were coded a zero for degree of brand loyalty toward the advertised brand. The three levels of the competitive position treatment were divided into two dummy variables to obtain a multiple correlation between competitive position and degree of brand loyalty toward the advertised brand. The multiple correlation coefficient is .33. This may result in downward biased F-values for the advertiser's competitive position.

Two interaction effects are significant: the interaction of advertisement copy type and theme and the interaction of competitive position and theme.

Table 2 presents the mean purchase intention ratings for the interaction of advertisement copy type and theme adjusted for the covariates. The lowest mean purchase intentions occur for the interaction of the non-comparative advertisement and theme two, while the highest purchase intention mean occurs for the interaction of the comparative advertisement and theme two. Thus, although there are no relative differences in the impact of comparative and non-comparative copy, the particular theme used with a comparative advertisement may make it more effective than a non-comparative advertisement. Both themes two and three produce higher purchase intentions with a comparative advertisement than with a non-comparative advertisement.

TABLE 2

Adjusted Mean Purchase Intentions for the Interaction of Advertisement Copy Type and Theme<sup>a</sup>

Treatment Level	Mean Rating	N
Non-Comparative Ad X Theme 1	3.68	101
Non-Comparative Ad X Theme 2	3.00	101
Non-Comparative Ad X Theme 3	3.34	97
Comparative Ad X Theme 1	3.47	96
Comparative Ad X Theme 2	3.97	100
Comparative Ad X Theme 3	3.44	99

<sup>a</sup>  $p \leq .05$

The adjusted mean purchase intention ratings for the interaction of the advertiser's competitive position and theme are reported in Table 3. The highest purchase intention ratings occur for the interaction of the number three brand and theme three, and the lowest purchase intention ratings occur for the interaction of the number one brand and theme three. Thus, different themes employed by the same advertiser may differently affect the influence of his advertisement. These interaction effects allude to the value of copy testing.

TABLE 3

Adjusted Mean Purchase Intentions for the Interaction of Advertisement Copy Type and Theme<sup>a</sup>

Treatment Level	Mean Rating	N
New Brand X Theme 1	3.70	70
New Brand X Theme 2	3.18	69
New Brand X Theme 3	3.17	67
Number One Brand X Theme 1	3.09	64
Number One Brand X Theme 2	4.04	66
Number One Brand X Theme 3	2.92	64
Number Three Brand X Theme 1	3.93	63
Number Three Brand X Theme 2	3.24	66
Number Three Brand X Theme 3	4.08	65

<sup>a</sup>  $p \leq .01$



The summary of the analysis of variance for product characteristic importance ratings is presented in Table 4. The results indicate a significant difference in mean importance ratings across respondents. The summary of mean importance ratings is presented in Table 5.

TABLE 4

Analysis of Variance for Product Characteristic Importance Ratings

Source	d.f.	Sum of Squares	F-Ratio
Mean	1	243,786.40	
Products	8	7,136.58	169.22 <sup>a</sup>
Error	5337	28,134.66	

<sup>a</sup>  $p \leq .01$

TABLE 5

Mean Importance Ratings for Product Characteristics

Characteristic	Mean
Grand Mean	6.75
Price	5.87
Scent	5.89
Stops odor	8.20
Stops wetness	7.31
Applies dry	4.65
Doesn't stain	7.16
Doesn't sting	5.88
Long-lasting	7.72
Over-all effectiveness	8.09

The two product characteristics for which copy comparisons were made between the advertised brand and the competing brand in the comparative advertisements are stops wetness and odor. The highest mean for the importance ratings occurs for the characteristic stops odor. Stops wetness has the fourth highest mean rating and is higher than the grand mean. The second and third highest mean occur for over-all effectiveness and long-lasting, respectively. Overall effectiveness was not mentioned in any copy treatments, but the characteristic long-lasting appeared in all copy treatments and was not used for comparisons. Except for the product characteristic does not stain, all other means are lower than the grand mean. Thus, in the aspects of the copy which were not comparative, both relatively important and relatively unimportant characteristics were used, while the characteristics used for comparison appear to be relatively important to the respondents. This feature of the copy should serve to strengthen the comparisons.

The purpose of the validation study was to provide a measure of external validity concerning the generalizations that could be made from copy manipulations to full advertisements in print media. The results of the analysis of covariance for the validation study are summarized in Table 6. The terms included in the model

explain twenty-one percent of the variation of purchase intention ratings. The probability that these results occurred by chance is less than one percent. The full advertisements used in the validation study produce no significant main effect of advertisement copy type. Thus, purchase intentions for the comparative advertisement do not differ significantly from those for the non-comparative advertisement. The results of the validation study are similar to the results presented in Table 1 for the variables in common.

TABLE 6

Validation Study: Analysis of Covariance for Purchase Intentions

Source	d.f.	Sums of Squares	F-Ratio
Regression <sup>b</sup>	3	102.04	6.39 <sup>a</sup>
Aloyal <sup>c</sup>	1	101.15	19.00 <sup>a</sup>
Cloyal <sup>d</sup>	1	2.36	.44
Ad	1	.03	.01
Error	73	388.76	

<sup>a</sup>  $p \leq .01$

<sup>b</sup>  $R^2 = .21$ ; Intercept = 3.28

<sup>c</sup> Degree of brand loyalty toward the advertised brand

<sup>d</sup> Degree of brand loyalty toward the competing brand

#### Conclusions and Implications

Comparison of the results from the full advertisements and the copy manipulations tends to indicate that consumers react to full advertisements in essentially the same manner as they react to copy manipulations. Thus, under the conditions of this study, conclusions can be drawn from the copy manipulations which may be generalized to represent consumer reactions to full advertisements.

When isolated from other variables, a comparative advertisement is no more effective an influence upon purchase intentions than is a non-comparative advertisement. The results of the product characteristic ratings indicate that the characteristics used for comparison in the comparative copy were important to the respondents. Thus, any lack of significant differences in comparative and non-comparative purchase intentions probably cannot be attributed to the lack of importance of product characteristics used for comparison. It appears that, given the conditions of this experiment, an advertiser is neither going to increase nor decrease purchase intentions relative to a non-comparative advertisement by using a comparative advertisement.

However, in conjunction with specific themes a comparative advertisement may have a relatively stronger influence upon purchase intentions than a non-comparative advertisement. Thus, the choice of an advertising strategy should be mediated by the specific themes under consideration. An advertiser cannot simply assume that a non-competitive advertisement will have no relative effect upon purchase intentions toward his product without considering the impact of his theme upon the effect of his advertising strategy.

The theme of the advertisement appears to be an important

influence upon the effects of other variables on purchase intentions. Not only does the theme temper the relative impact of a comparative advertisement, but the theme also interacts with the advertiser's competitive position to differentially influence purchase intentions. Thus, certain themes may be more effective for advertisers in a particular competitive position than are other themes. An advertiser should consider his market position as unique and use the theme which most favorably affects the impact of his competitive position upon consumer response.

Test results appearing in advertisements are receiving scrutiny by public policy makers and major television networks due to the possibility that deceptive claims may mislead consumer expectations and purchasing. However, the results of this study tend to indicate that substantiated claims are relatively no more effective in influencing purchase intention ratings than are unsubstantiated claims. This does not mean that public policy makers and other groups should not be concerned about the impact of deceptive claims or test results. The statements for claim substantiation only referred to an independent testing agency and not to a specifically named agency. Possibly, the inclusion of a named agency would make the claim substantiation stronger to the extent that it would be more effective than unsubstantiated claims. In addition, repeated exposure to an advertisement may produce differences in the effects of different strategies incorporated into this study. Finally, since consumers have been so frequently exposed to claim substantiation, inoculation may have occurred, and the assumptions surrounding the effectiveness of claim substantiation may not be valid.

No significant main effects occur for any of the uncontrolled independent variables. Overall, the lack of significant main effects and the presence of significant interactions tends to imply that the advertising industry is operating in a multivariate world. The effects of the variables in this study do not constitute a "simple" situation, and "theories" or assumptions based on main effects only are not rich and do not adequately describe the situation. Basically, what this research implies is that general statements may be made concerning comparative advertisements only as indicated by the recurring underlying conditions and contexts.

Advertisers have been using comparative advertising and public policy makers, as well as other groups, have been endorsing it without empirical evidence regarding its effects upon the consumer. The endorsements are based upon the assumption that comparative advertising supplies the consumer with more information and more useful information than non-comparative advertising. The accuracy of these assumptions may have been alluded to by increased purchase intentions for comparative advertising. This was not the case. More information is needed concerning the impact of comparative advertising with regard to the assumptions and controversy surrounding the phenomenon. Future research should investigate the effects of comparative advertising upon other dependent variables for other products over-time.

#### References

A. Barr and J. Goodnight, A User's Guide to the Statistical Analysis System (Raleigh: North Carolina University, 1972), 114.

Tom Bradshaw, "Comparative Ads: What's Their Status Now?" Television/Radio Age (April 29, 1974).

Maurene Christopher, "NBC Spells Out New Formal Guide

lines for Comparative Spots," Advertising Age (January 28, 1974), 1.

W. Dixon, Biomedical Computer Programs (2nd ed.) (Berkeley: University of California Press, 1970), 586-600.

Don Grant, "Be Hard on Comparison Ads: Roberts to Four A's," Advertising Age (November 19, 1973).

Jack Roberts, "Comparative Advertising...I'm O.K... You're Not O.K.," Speech delivered to American Association of Advertising Agencies (New York: November 15, 1973).

George Szybillo and Richard Heslin, "The Television Code," (Washington, D.C.: National Association of Broadcasters, April, 1973).

S. Tannenbaum, "Policy Statement and Guidelines for Comparative Advertising," American Association of Advertising Agencies (April, 1974).

Stanley Ulanoff, "No Names Please: The Case for Comparative Advertising," Paper presented at the Annual Meeting of the American Academy of Advertising (University of Oklahoma: March. 1972).

## WHAT IS CONSUMERISM?

Kent L. Granzin, University of Utah  
Gary M. Grikscheit, University of Utah

### Abstract

This research is a response to the question: "What is consumerism?" The hypothesis that general attitudes toward the socio-economic environment influence more specific attitudes toward buyer-seller relations is evaluated using factor and canonical correlation analysis. Consumerism exhibits a complex structure consisting of alternative views on consumer protection. The findings provide a foundation for designing preventative and remedial programs.

### Introduction

Consumerism, the "social movement seeking to augment the rights and power of buyers in relation to sellers," (Kotler, 1972) is manifest in new laws, regulations, and marketing practices, as well as in new public attitudes toward government and business. Despite unprecedented affluence and a host of new and improved products and services, consumers are disenchanted with the marketing system. Symptomatic of consumers' malaise are complaints about rising prices, demands for improved products and services, suggestions for improving the adequacy of product information, and concern for the physical environment. In response to these consumer dissatisfactions, self-appointed spokesmen, unions, professional consumer organizations, business leaders, and politicians have advocated numerous remedies for curing consumer ills. During the rise of consumerism at least one consumer spokesman, Ralph Nader, has become a modern folk-hero.

Yet, as is typical for youthful social movements in a state of rapid change, the structure of consumerism is neither well defined nor clearly understood. To be sure, both participants and observers have expressed their perceptions of consumerism, but little has been done to empirically characterize this social movement. Hence, the purpose of this research is to provide at least partial answers to the question "What is Consumerism?"

Guiding the search for answers to the question is the hypothesis that general attitudes toward the socio-economic environment influence more specific attitudes toward buyer-seller relations. This hypothesis follows from the results of a number of studies reported in the behavioral literature. For example, Bolton's study of peace-group membership shows that a person's predisposition toward social action is related to his degree of alienation (Bolton, 1972). Zygmunt asserts the socio-cultural context is central to understanding movement affiliation, Longman and Pruden find anomie important to understanding attitudes toward specific marketing acts, and a number of other investigators report the viability of employing general measures to determine views on particular subjects (Zygmunt, 1972; Longman and Pruden, 1972; McCloskey and Schaar, 1965; Srole, 1956). Based on the behavioral and business literature, the present research employs three types of such general attitudinal measures to provide data for testing its hypothesis: (1) degree of reliance on business and/or government for regulating economic activities; (2) satisfaction with one's social and economic environment; and (3) trust of others.

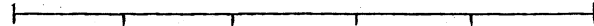
### Method

Based on a cluster sampling scheme, Salt Lake City was divided into 49 cells according to census tracts, and blocks within each tract were selected at random.<sup>1</sup> For each block a map detailing the number and location of dwelling units was drawn and the units to be contacted were selected on a systematic random basis. The sampling plan sought to include 20 percent of the dwelling units per block, alternating male and female respondents of at least 18 years of age. One callback was made before the next dwelling was substituted.

A total of 295 residents provided usable responses to the personally administered questionnaire. The instrument contained 28 opinion items representing both the three classes of general measures introduced above and items judged a priori to indicate respondents' specific interest in consumerism. All opinion items were 100 mm Likert-type scales such as:

"More protection is needed for the consumers of this country."

Strongly Agree	Slightly Agree	Slightly Disagree	Strongly Disagree
-------------------	-------------------	----------------------	----------------------



To avoid pattern responses by respondents, the instrument used some items stated as positive and some stated as negative opinions, an alternative approach to reversing the scales themselves. Naturally, the number of items that can reasonably be included in a questionnaire is limited by the willingness of respondents to cooperate with interviewers.

Business writings provide a rich source of the items expressing opinions on specific consumerism issues (Day and Aaker, 1970; Bauer and Greyser, 1967; Buskirk and Rothe, 1970). Typical of these measures is the item tapping support for the leaders of the consumerism movement: "Consumer spokesmen (like Ralph Nader) should be thanked for the work they are doing."

A number of more general items involve support for business and/or government control of marketing operations, based on the rationale that consumerism frequently calls on either or both of these institutions to protect the

<sup>1</sup>The reader may be interested in speculating about how findings are affected by local demographic characteristics even though this research does not attempt to estimate parameters for a universe larger than residents of the Salt Lake metropolitan area.

According to the 1970 census, Salt Lake City, with a population of 175,798, was 98.1 percent whites and 1.9 percent nonwhite. The Spanish-American ethnic group (primarily recorded as white) constituted 4.8 percent of the total, while persons of foreign stock contributed 15.4 percent of the total population. Also of possible relevance, although local scholarly studies suggest otherwise, is the predominance of a single religion. Approximately 47 percent of the Salt Lake residents claim affiliation with the L.D.S. (Mormon) Church.

consumer (Herman, 1970). Underlying their inclusion was the desire to see whether the consumer views the two institutions as complementary or alternative means for furthering his ends. That is, are business and government perceived as equally responsible for remedying wrongs? Direct remedy by these institutions is exemplified by the items "Business should bear more of the cost of consumer protection" and "The government should do more to control prices." Government control of the economy varies, however, in the directness of its link to business matters. While control of price, product, and promotion matters reaches to the heart of business strategy formulation, government is also involved with more indirect issues. To learn the extent to which consumerism includes or overlaps with a desire for greater control of society in a larger sense, this research includes such items as that asking consumers to what degree they agree that "The government should do more to regulate and control TV programming."

To measure satisfaction and trust, the relevant measures for one's general frame of reference are again limitless. However, the literature of the behavioral sciences provides strong guidance (Clark, 1959; Gerson, 1965; McCloskey and Schaar, 1965). Typical of items used for measuring general satisfaction is "It would be accurate to say this country is going to the dogs." Trustfulness is expressed by measures like "The average person you meet is honest."

To this point discussion has considered a categoric and necessarily conceptual dichotomization of items into specific consumerism interests and general attitudes toward the socio-economic environment. This characterization is supportable and necessary for data collection, but somewhat arbitrary. To avoid an arbitrary separation during analysis the research required a means for empirically isolating specific consumerism items as seen by the respondents. Therefore, factor analysis was employed to separate the 28 items not only into specific and general sets, but to divide the latter category into its constituent subsets. Factor analysis by the principal components method was employed for this task, with orthogonal rotation of factors by the varimax procedure (Kaiser, 1958). Six factors were rotated toward simple structure, using the criterion for selection that this number (lowest eigenvalue 1.16) provided maximum clarity of factor structure for the analysis to follow (Hakstian and Muller, 1973; John W. Thompson, 1962).

Having isolated the single factor most closely representing specific consumerism interest, the analysis proceeded to examine the relationships between this set of consumerism variables and the more general items. Canonical correlation analysis made it possible to examine simultaneously the multivariate relationships between specific and general measures. Canonical correlation extends multiple correlation to include a set of variables as criteria, rather than using only a single criterion variable (Stewart and Love, 1968). As in multiple correlation, the predictor set also involves multiple variables. The resulting equations are linear functions which portray the relationships between the two sets of variables. Thus, the functions developed here relate the set of specific consumerism variables to the set of more general predictor variables in a way such that the sets are correlated maximally with each other. Correlations between the best weighted combination of criteria and best weighted index of predictors are computed so that each variate is orthogonal to those already computed for earlier functions representing its set. Examining the correlation between each variable and its canonical variate is analogous to examining factor loadings (Veldman, 1967).

The analysis computes as many of these linear functions as there are variables in the smaller of the two sets, in this case six.

## Results and Discussion

Table 1 presents the factor structure of the 28 opinion items. The following interpretation of this structure considers each of the six factors in turn, giving greatest emphasis to those variables whose loadings on primary and secondary factors differ by at least .1, indicating they are relatively unambiguous (Neulinger and Breit, 1969).

Factor I--Trust in Americans. Central to this factor is the opinion that the average person is honest. Manufacturers may be personified by consumers and these producers' consideration of their customers associated with a single concerned individual representing the firm. Responses showing support for the American culture sharpen the picture of a basic dimension indicating a need (or lack of need) for protection from other residents of the respondent's country. This factor reflects more personalized consideration of other individual Americans than the factors that follow.

Factor II--Distrust and Dissatisfaction with Institutions. Factor II, which may be interpreted as distrust/dissatisfaction (or trust/satisfaction) appears less personalized than Factor I. Dissatisfaction expressed by agreement that the country is going to the dogs, the economy is in a sad state, and the government is not doing enough for peace shifts the focus to the national social and economic scene, indicating concern (or lack of concern) with the way the federal government is doing its job. Similarly, distrust of government officials, military leaders, and government leaders appears to indicate a concern with institutions and the faceless officials who are responsible for their operation.

Factor III--Need for Environmental Change. Both willingness to believe Russian leaders on peace and desire for increased government action on pollution problems indicate interest in changing the current situation. Unfortunately, only two items load on Factor III, making interpretation difficult. Nonetheless, the fact that this is the only factor with no items directly linked to business suggests the broad environmental nature of the perceived need for change underlying this dimension.

Factor IV--Need for Institutional Change. Factor IV combines two related aspects. One is the broad concern for the current situation revealed as agreement on need for improving race relations and the country's educational situation. In this country responsibility for such improvement is currently delegated to government agencies. The second concern is the need for business to take a greater responsibility to regulate its conduct and protect the consumer. Thus, the two institutions of government and business are here called upon to improve their performance.

Factor V--Need for Increased Consumer Protection. Like the preceding two factors, Factor V indicates a need for changes directly related to new protective laws, organization of consumers, rewards for those advocating change, and explicitly, more consumer protection. Two ambiguous items are also included. The belief that advertising is misleading offers one rationale for change, and the call for stronger penalties for cheating firms furnishes a means for bringing about increased protection.

Factor VI--Call for Increased Governmental Regulation. Preceding factors have expressed the underlying reason for and the need for change in the conduct of the economic system. In addition, Factor IV assigns that responsibility to business firms. On the other hand, Factor VI calls for government to exercise the necessary control. Among the unambiguous items, government is given sanction to control prices, regulate salesmen and sales practices, and test, rate, and/or grade certain consumer products. Power to regulate is extended to TV programming, a product of the television industry, and greater power is desired for the government agencies involved in regulating business. Again, two ambiguous items load highest on this factor. Approval of the government's role in promoting auto safety and granting government more power to promote racial integration also overlap the concerns of other factors, as shown by their secondary loadings.

Two criteria provide powerful support for selecting Factor V as that most central to the concept of specific interest in consumerism. First, the word "consumer" appears in four of the six items loading highest on the factor. Second, an analysis at the conceptual level confirms Factor V as the basic dimension most closely tied to the call for augmented "rights and powers of buyers." For these reasons, Factor V appears to most closely represent the heart of the consumerism

movement, and this factor and its associated six variables are selected as the criterion set for the canonical analysis.

Nonetheless, it must be recognized that, as expected, other items expressive of the consumerism movement load most highly on other factors, especially Factors IV and VI. Thus, while the picture of the heart of consumerism is presented in basic form through Factor V, the responsibilities of both business and government in providing for these needs (Factors IV and VI) remain to be clearly determined. The next phase of the analysis investigates these interactions, and also adds the supplementary explanatory power of items loading on Factors I, II, and III as well.

Based on the preceding results, canonical analysis provides six canonical functions representing correlations between the best weighted combination of the six variables loading highest on Factor V and a similar linear combination of the remaining 22 items. Five of the six linear functions have correlations significant at the .05 level as indicated by  $X^2$  tests of their respective Wilks' lambdas. Probabilities of a Type I error for Functions A through E are .000, .000, .022, .032, and .040 respectively. Function F is non-significant. In addition, Function E, while significant, has such low loadings that its interpretation would add

TABLE 1  
FACTOR STRUCTURE OF 28 OPINION ITEMS CONCERNING  
CONSUMERISM AND RELATED ISSUES

Composition of Factors	Factor Loadings*					
	I	II	III	IV	V	VI
<u>I Trust in Americans</u>						
Manufacturers consider their customers	76	-18	04	-07	-10	-03
Average person is honest	71	-08	04	08	06	-08
Rather live in America than elsewhere	55	-26	-38	03	09	-01
<u>II Distrust and Dissatisfaction with Institutions</u>						
This country is going to the dogs	-30	65	-05	-02	-14	23
Hard to trust government officials	-24	64	04	15	25	-08
Military leaders not to be trusted	-15	63	33	01	17	-08
Firms market dangerous products	-21	60	-03	-07	19	18
Economy is in a sad state	01	57	-18	38	12	22
Government not doing enough for peace	15	53	22	21	05	02
<u>III Need for Environmental Change</u>						
Should believe Russian leaders on peace	01	07	70	11	04	10
Govt do more on pollution control	-19	-05	47	24	32	33
<u>IV Need for Institutional Change</u>						
Business should regulate its conduct more	01	-15	03	66	21	10
Concerned about today's race relations	08	31	22	57	-07	07
Business bear more cost of consumer protection	-03	01	05	53	21	31
Educational system needs improving	-07	33	06	52	07	-10
<u>V Need for Increased Consumer Protection</u>						
Consumer spokesmen should be thanked	-01	-10	25	05	70	03
Need some way to organize consumers	06	22	-01	-03	62	32
Need new laws to protect consumers	-03	24	02	24	61	13
More protection needed for consumers	10	11	01	30	53	15
Advertising tends to be misleading**	-14	39	-13	04	48	06
Need stronger penalties for cheating firms**	05	18	-16	36	37	17
<u>VI Call for Increased Governmental Regulation</u>						
Govt do more to control prices	-10	19	-11	10	15	69
Govt do more to regulate salesmen	-22	03	05	18	08	68
Govt agencies have more power to regulate	-15	03	20	09	04	65
Govt do more to regulate TV programming	14	-03	06	-06	06	63
Govt do more to promote racial integration**	16	24	51	-12	00	58
Govt test, rate, grade consumer products	05	12	-05	33	29	54
Glad govt improves auto safety**	02	-15	14	31	36	45

\* Decimals omitted for clarity; all values in hundredths

\*\* Ambiguous variables with primary and secondary loadings differing by <.1

little to understanding the consumerism issues involved.

Discussion of Function A considers only correlations between items and canonical variates of .5 or greater. Criterion items on need for more protection, organizing consumers, new protective laws, and stronger penalties for cheating firms, as well as feeling advertising is misleading, relate positively to six other opinion items. Three of these other measures come from Factor II, Distrust and Dissatisfaction with Institutions: distrust of government officials, fear that firms market dangerous products, and feeling the economy is in a sad state. This canonical relationship provides solid evidence for the hypothesized link between a general societal malaise and specific remedies to bring more protection for the individual consumer. An item from Factor IV, the call for business to bear more of the cost of consumer protection, also enters positively. While the respondents whose opinions are described by this function show distrust of government officials, they do see a need for more government regulation, as shown by the two Factor VI items calling for control of prices and testing of consumer products. In short, Function A describes a general discouragement with one's situation. In particular, the persons described call for increased consumer protection. They want business to bear its cost and government to provide the necessary control.

In function B four items have loadings over .4. Thanking consumer spokesmen, absent from the first equation, provides the only consumer protection item here. This criterion item correlates positively with predictors expressing need for firmer government action on pollution control and more self-regulation by business. However, its negative correlation with the idea that firms market dangerous products underscores the more specific nature of this relationship as compared with Function A. In this function, appreciation of consumer spokesmen relates to their championing a move for more corporate responsibility in areas of environmental interest, as opposed to concern for the price, product, and promotion aspects of marketing strategy.

Examination of loadings of .4 and over for Function C shows that a segment of consumers is not concerned about misleading advertising, but wants more protection for consumers. The relationship indicates trust (a lack of distrust) of government officials and agreement there is need for price control. Apparently some respondents see the need for a particular remedy, price control. This temporal issue, like that underlying the environmental concern expressed by Function B, again indicates a specific concern. This more specific interest appears more reasoned than the general interest shown in A, and the general distrust and dissatisfaction concerns do not enter here.

TABLE 2  
CANONICAL LOADINGS FOR TWO SETS OF OPINION ITEMS  
CONCERNING CONSUMERISM AND RELATED ISSUES

Opinion Items	Loadings on Six Sets of Canonical Variates <sup>a</sup>					
	A	B	C	D	E	F
<u>Need for Increased Consumer Protection Items</u>						
More protection needed for consumers	59 <sup>b</sup>	31	40 <sup>b</sup>	-30	39	17
Need some way to organize consumers	62 <sup>b</sup>	-06	36	41 <sup>b</sup>	-16	63
Consumer spokesmen should be thanked	36	72 <sup>b</sup>	-38	14	17	25
Advertising tends to be misleading	68 <sup>b</sup>	-35	-58 <sup>b</sup>	-40 <sup>b</sup>	24	16
Need new laws to protect consumers	69 <sup>b</sup>	32	03	-14	-57	02
Need stronger penalties for cheating firms	68 <sup>b</sup>	04	22	40 <sup>b</sup>	00	-60
<u>All Other Opinion Items</u>						
Manufacturers consider their customers	-23	08	18	20	12	-06
Average person is honest	-02	33	-01	-04	22	-10
Rather live in America than elsewhere	-04	09	25	17	06	-31
This country is going to the dogs	25	-34	-23	-01	-28	-03
Hard to trust government officials	53 <sup>b</sup>	-03	-44 <sup>b</sup>	-32	-28	18
Military leaders not to be trusted	36	08	-32	-20	-18	-03
Firms market dangerous products	56 <sup>b</sup>	-49 <sup>b</sup>	-33	-16	11	09
Economy is in a sad state	60 <sup>b</sup>	-08	00	-02	-17	05
Government not doing enough for peace	32	06	19	-03	11	24
Should believe Russian leaders on peace	19	14	-04	-06	08	15
Govt do more on pollution control	46	43 <sup>b</sup>	04	18	-28	-01
Business should regulate its conduct more	36	51 <sup>b</sup>	08	-22	12	02
Concerned about today's race relations	38	01	24	-13	-14	-06
Business bear more cost of consumer protection	52 <sup>b</sup>	24	-01	14	17	-16
Educational system needs improving	42	04	03	-26	24	-43
Govt do more to control prices	58 <sup>b</sup>	-24	43 <sup>b</sup>	31	-11	15
Govt do more to regulate salesmen	39	07	09	-20	21	40
Govt agencies have more power to regulate	27	09	22	18	-27	41
Govt do more to regulate TV programming	17	20	20	24	-04	14
Govt do more to promote racial integration	20	06	31	28	-22	51
Govt test, rate, grade consumer products	63 <sup>b</sup>	12	35	29	-19	15
Glad govt improves auto safety	46	39	00	47 <sup>b</sup>	18	20

<sup>a</sup> Decimals omitted for clarity; all values in hundredths

<sup>b</sup> Indicates loadings for a given canonical root considered in discussion

Function D also has four loadings of .4 or more. Three of these items are criterion variables and represent perceived need for consumer aid: agreement on the need for more consumer organization and stronger penalties for cheating firms, but disagreement that advertising tends to mislead. The single predictor item comes from Factor VI, approval of the government's stepping in on auto safety. Thus, this relationship supports a specific government activity and would apparently back strong government action to exact compliance from the auto makers. Rather surprisingly, this relationship indicates a desire for consumer organization but does not to a marked degree care to reward consumer spokesmen. Considering the fame accruing to Ralph Nader for his attacks on the auto industry, one may wonder just how pervasive is his support. The moderate negative loading on distrust of government officials solidifies this direct appeal for government action, quite possibly as a preferred alternative to endorsing the leadership of consumer spokesmen.

In sum, the canonical functions examined show both broad and narrow aspects of consumerism. The strongest relationship exhibits a link between the call for more consumer protection and a general distrust in and dissatisfaction with the current performance of America's institutions. Here, both business and government are called upon to act in the interests of the consumer. But superimposed on this strong general correlation are three lesser specific relationships. In particular, pollution control, price control, and auto safety emerge as issues of importance to a number of those responding to the survey.

#### Conclusion

Consumerism emerges as a complex cluster of issues. However, most salient appears to be the concern for consumer protection. Peripheral are concerns for how this protection may be gained. On one hand, business may bear its cost. On the other, government regulation may be the consumer's salvation. Significantly, these basic positions are alternative rather than complementary.

The desire for increased consumer protection reflects something other than a studied appraisal of the workings of America's economic system. For many citizens, their dissatisfaction and discouragement with the operation of society appears to bring a general call for more consumer protection. For remedies, many of these dissatisfied persons may choose government control as possibly the best alternative means of correction from a set of undesirable remedies. For other citizens the issues are more clear-cut and they view pollution control, regulation of prices, and auto safety as specific concerns. Had other such items been included, they too might have emerged as specific concerns.

Business leaders and government officials concerned with bringing about change would do well to consider the structure of consumerism presented here. However, they should recognize that the total picture must be balanced by considering the citizens whose opinions form the opposite ends of the relationships expressed here. While some respondents showed dissatisfaction, distrust and need for redress, others indicated satisfaction, trust, and approval of the current situation. Thus, a survey of national scope would be useful for assessing the relative weights to be given to the positions revealed in this report. Nonetheless, the empirical knowledge of the basic structure of consumerism presented here provides an important foundation for formulating preventative and remedial action.

We gratefully acknowledge permission from the Journal of Business Research to reprint portions of this article from "Who are the Consumerists?" by Gary M. Grikscheit and Kent L. Granzin. © 1975 by the Journal of Business Research.

#### References

- Raymond A. Bauer and Stephen A. Greyser, "The Dialogue That Never Happens," Harvard Business Review, 45(November/December, 1967),186-190.
- Charles D. Bolton, "Alienation and Action: A Study of Peace-Group Members," American Journal of Sociology, 78(November, 1972),537-561.
- Richard H. Buskirk and James T. Rothe, "Consumerism--An Interpretation," Journal of Marketing, 34(October, 1970),61-65.
- John P. Clark, "Measuring Alienation Within A Social System," American Sociological Review, 24(December, 1959),783-791.
- George S. Day and David A. Aaker, "A Guide to Consumerism," Journal of Marketing, 34(July, 1970),12-19.
- Walter M. Gerson, "Alienation in Mass Society: Some Causes and Responses," Sociology and Social Research, 49(January, 1965),143-151.
- A. Ralph Hakstian and Victor J. Muller, "Some Notes on the Number of Factors Problem," Multivariate Behavioral Research, 8(October, 1973),461-475.
- Robert O. Herrman, "Consumerism: Its Goals, Organizations and Future," Journal of Marketing, 34(October, 1970),55-60.
- Henry F. Kaiser, "The Varimax Criterion for Analytical Rotation in Factor Analysis," Psychometrika, 23(September, 1958),187-200.
- Philip Kotler, "What Consumerism Means for Marketers," Harvard Business Review, 50(May/June, 1972),48-57.
- Douglas S. Longman and Henry O. Pruden, "Alienation from the Marketplace: A Study in Black, Brown, and White," in Combined Proceedings, 1971 Spring and Fall Conferences, ed. by Fred C. Allvine (Chicago, Illinois: American Marketing Association, 1972).
- Herbert McCloskey and John H. Schaar, "Psychological Dimensions of Anomie," American Sociological Review, 30(February, 1965),14-40.
- John Neulinger and Miranda Breit, "Attitude Dimensions of Leisure," Journal of Leisure Research, 1(Summer, 1969),108-115.
- Leo Srole, "Social Integration and Certain Corollaries: An Exploratory Study," American Sociological Review, 21(December, 1956),709-716.
- Douglas Stewart and William Love, "A General Canonical Correlation Index," Psychological Bulletin, 70(August, 1968),160-163.
- John W. Thompson, "Meaningful and Unmeaningful Rotation of Factors," Psychological Bulletin, 59(May, 1962), 211-223.
- Donald J. Veldman, FORTRAN Programming for the Behavioral Sciences, (New York: Holt, Rinehart, and Winston, 1967).
- Joseph F. Zygmunt, "Movements and Motive: Some Unresolved Issues in the Psychology of Social Movements," Human Relations, 25(November, 1972),449-467.

## CONSUMER ACTIVISTS: WHAT MAKES THEM DIFFERENT?

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

### Abstract

Throughout the debate on the consumer movement, few attempts have been made to determine the characteristics of the leaders of the movement or to differentiate the vocal consumer from the passive. For the most part, the research which has been done in this area has examined the characteristics of those people who say they are interested in consumer matters. Using a behavioral measure, this paper will develop a profile of consumers who have become involved in the consumer movement. The question will be asked and answered: "What makes the consumer activist different from the average consumer?"

### Introduction

Much has been written in recent years about the growth and activities of the "consumer movement" and about the direction in which "consumerism" will move in the future. It has been generally acknowledged that consumerism is but an outward and early manifestation of an underlying social concern which has been sweeping our society. The current consumer movement, which developed during the 1960's, is now part of a more general concern for the quality of the environment, the rights of individuals, protection of privacy, the state of the economy, the impersonality of big business and big government, and other related social and economic issues.

Throughout the debate on the consumerism phenomenon of the 1960's and 1970's, few efforts have been made to determine the characteristics of the leaders of the consumer movement or to differentiate the vocal consumer from the passive. This paper will develop a profile of consumers who have become involved in the consumer movement and should prove to be of interest to marketers and public policy makers for a number of reasons: (1) it marks the first time an attempt has been made to profile consumers who have joined a consumer association; (2) it will determine for the marketer whether the consumer activist constitutes a sufficiently important and viable market segment to warrant development of marketing programs designed to appeal to consumers who exhibit the characteristics of the activist; (3) it will examine the consumer activist in detail from a number of different perspectives; (4) it will determine those characteristics which best distinguish the consumerist from the mass of consumers who are more passive on consumerism and other social issues.

### Literature Review

Previous attempts to examine the characteristics of the consumer activist have suffered from a number of weaknesses which cast doubt on their appropriateness for the development of marketing programs and government programs in the area of consumer affairs.

#### Weaknesses

Firstly, authors in this area have generally relied upon surrogate measures of consumer activism, assuming

that consumers who scored high on a Social Responsibility Scale (Anderson and Cunningham, 1972), or who expressed a concern over environmental pollution (Kinnear and Taylor, 1973), or who expressed negative attitudes toward business (Hustad and Pessemier, 1973) are consumer activists.

Secondly, research dealing with consumer activism has, almost without exception, relied upon psychological measures of the construct.

Thirdly, most studies have relied simply upon demographic and socio-economic variables to profile the consumer activist.

Fourthly, with the exception of the work of Kinnear and Taylor (1973), previous studies have not incorporated Canadian data, with the result that little work has been done to examine the consumer activist in this country.

### Findings

While these previous studies have not employed a measure of actual consumer activism, some of their findings are, nonetheless, interesting. Hustad and Pessemier (1973) found that subjects who were strongly anti-business were higher in employment status and education level. Barksdale and Darden (1972) found that younger, more liberal respondents were more critical of marketing and more impressed with the accomplishments of the consumer movement. The Roper Organization (1971) found younger, more affluent consumers to be more concerned with environmental pollution and more willing to contribute to its reduction. Kinnear, Taylor and Ahmed (1974) found ecologically concerned consumers to be high in income and to be more open to new ideas, and higher in the need to satisfy intellectual curiosity and the need to obtain personal safety. Anderson and Cunningham (1972) assumed that subjects who scored high on the Social Responsibility Scale (SRS) would also manifest social consciousness in consumption decisions. Those who scored high on the SRS were younger and characterized by higher occupational and socio-economic levels. The high scorers on the SRS were also less alienated, less conservative, and more cosmopolitan than subjects in the low socially responsible group.

These findings have relevance for the study at hand only insofar as the dependent measures employed in these previous studies are assumed to be correlated highly with consumer activism.

### Research Methodology

#### The Sample

The sample consisted of two groups: the first contained members of the Market Facts of Canada Consumer Mail Panel (MFP) resident in the Toronto Metropolitan Census Area; the second was comprised of members of the Consumers' Association of Canada resident in the same area. MFP members were found to be quite representative of the population at large (Tigert, Barnes and Bourgeois, 1975). Membership in the CAC represents a behavioral measure of consumer activism.



Questionnaires were mailed to 1327 MFP members and to 600 CAC members. These mailings yielded 968 completed, usable questionnaires from the MFP group and 267 from the CAC group. Both samples were restricted to female heads of households because of the importance of the wife and mother in the family buying decision and because the membership of the CAC is predominantly female.

#### The Questionnaire

The questionnaire employed in this study consisted of three main sections. A fourth section containing demographic and socio-economic questions was added to the questionnaire for mailing to CAC members. These data were provided on MFP members by the panel operator.

The first section of the questionnaire consisted of 139 "life style" statements to which subjects were asked to indicate their level of agreement on a six-point, Likert-type scale. Each of the statements in this first section pertained to subjects' activities, interests and opinions in the areas of retail shopping and the mass media. The second section contained questions pertaining to subjects' exposure to and attitudes toward the mass media and selected retail stores. The final section of the questionnaire contained two retail newspaper advertisements to which subjects were asked to indicate their reactions by responding to twenty-one "life style" statements, each of which pertained to one of the following elements of reaction: 1. perceived believability; 2. perceived value for the money; 3. motivation to initiate action leading to purchase.

#### The Data

Blanks in the data, where subjects were found not to have answered particular questions, were replaced with the mean score on those questions for the group (MFP or CAC) to which the subject belonged. The effect of this procedure was to provide complete data for all subjects in the sample, a necessary condition for most analytical procedures.

The large number of variables in the data base was reduced through the use of principal components factor analysis. This approach to data reduction was employed in four sections of the questionnaire and was used to replace subjects' scores on a large number of variables with their sum scores on factors which represent underlying dimensions of subjects' attitudes or behavior. Responses to "life style" statements in the first and third sections of the questionnaire were factor analysed and sum scores calculated for subjects on factors which represent underlying attitudes toward retail shopping, the mass media, and reaction to advertising. Similarly, information on readership and listenership of a number of individual magazines and radio stations was factor analysed and produced factors which represent readership of types of magazines and listenership of types of radio stations.

#### Analysis

The univariate analysis of the data involved a three-step process. Firstly, the mean scores on each variable were calculated for each of the two groups of subjects. Secondly, t-tests were performed on each variable to test the significance of the difference between the mean scores of the MFP and CAC groups. Thirdly, those variables which produced a significant t-statistic in comparing the group means were examined in a cross-tabulation analysis.

#### Demographic and Socioeconomic Variables

Results obtained from comparing MFP and CAC subjects on these variables are presented in Table 1 and are found to be quite interesting and support some of the findings reported in earlier studies. The consumer activist was found to have attained a much higher education level than had the mail panel member. The cross-tabulation analysis revealed that 67.8% of CAC members had attended college as compared with 22.5% of MFP members. The results, when comparing the groups on husband's education level, are equally striking. Similarly, CAC members are more likely to be employed outside the home than are MFP members and, when employed, are more likely to hold professional, management, or supervisory positions. The husbands of CAC members are also more likely to hold such positions. One result of such differences is that fewer CAC members are found in lower income groups.

TABLE 1

UNIVARIATE ANALYSES OF DEMOGRAPHIC AND SOCIOECONOMIC VARIABLES

VARIABLE	T-TEST	SIG. LEVEL	CHI-SQUARE	SIG. LEVEL
Type of home	-4.388	0.001	17.371	0.001
Home ownership	-2.079	0.05	9.826	0.002
Respondent education	-13.359	0.001	195.665	0.001
Husband education	-11.674	0.001	89.537	0.001
Employment Status	3.205	0.01	9.995	0.007
Respondent occupation	21.593	0.001	70.513	0.001
Husband occupation	8.638	0.001	20.642	0.001
Family income	-5.959	0.001	23.664	0.001
Age	-1.460	>0.05	*n.c.	
Household size	6.791	0.001	47.896	0.001

\*the cross-tabulation analysis was not performed and the chi-square statistic not calculated for those variables which failed to produce a t-statistic which was significant at the 0.05 level.

It was found that consumer activists are more likely to live in apartments and townhouses than are MFP subjects and are less likely to own their own homes. In addition, CAC members appear to reside in households with fewer members. These results may be explained, in part, by the fact that the mail panel understates the number of unmarried female heads of households who reside in apartments (Tigert, Barnes and Bourgeois, 1975).

#### Life Style Variables

Table 2 presents a summary of the univariate analysis on the "life style" variables. Fourteen life style variables produced differences between groups which were significant at  $\alpha < 0.05$ . Significant mean differences were found between the MFP and CAC groups on six life style factors. CAC members generally expressed less interest than did MFP members in shopping for bargains and also tended to derive significantly less enjoyment from shopping. The consumer activist also expressed more negative attitudes toward shopping at discount houses. In media-related matters, the CAC member was found to be more skeptical of the accuracy of news reporting by the mass media and expressed much more negative attitudes toward media advertising than did members of the mail panel. The negative attitude of the consumer activist toward television as an information source was also revealed in this section of the analysis and is discussed in more detail later in this paper.

TABLE 2

## UNIVARIATE ANALYSES OF 'LIFE STYLE' VARIABLES

VARIABLE**	T-TEST	SIGNIFICANCE LEVEL	CHI-SQUARE	SIGNIFICANCE LEVEL
'bargain interest'	7.141	0.001	43.701	0.001
'confidence in the mass media'	-5.269	0.001	17.627	0.001
'reliance on newspapers as a source of information'	-0.853	>0.05	*n.c.	
'enjoyment of shopping'	7.788	0.001	37.061	0.001
'reliance on others for shopping information'	1.931	>0.05	n.c.	
'confidence in shopping at department stores'	1.927	>0.05	n.c.	
'attitude toward the use of credit'	0.173	>0.05	n.c.	
'confidence in television as a source of information'	6.244	0.001	43.334	0.001
'confidence in shopping at discount stores'	5.449	0.001	25.364	0.001
'skepticism regarding retail sales'	1.665	>0.05	n.c.	
'confidence in private label products'	0.002	>0.05	n.c.	
'confidence in broadcast media advertising'	7.271	0.001	65.102	0.001
"I like sales clerks to leave me alone until I find what I want to buy"	4.410	0.001	20.419	0.001
"I like to shop alone"	-5.285	0.001	16.405	0.001
"I watch a lot of programs on educational TV"	-4.649	0.001	16.127	0.001
"I spend a lot of time talking to friends about products and brands"	6.068	0.001	26.460	0.001
"I often spend more money than I had planned when shopping"	2.330	0.05	3.725	0.156
"Price is a good indication of the quality of a product"	-1.357	>0.05	n.c.	
"Too much television is bad for children"	-5.981	0.001	22.783	0.001
"When I see a new brand on the shelf, I often buy it to see what it is like"	1.303	>0.05	n.c.	
"To me, it is more important for clothing to last several seasons than it is for it to be the latest style"	-5.832	0.001	17.298	0.001
"To me, price is the most important thing to consider when shopping"	5.656	0.001	20.882	0.001

\* the crosstabulation analysis was not performed and the chi-square statistic not calculated for those variables which failed to produce a t-statistic which was significant at the 0.05 level.

\*\*the first twelve variables represent factors which were obtained through factor analysis of subjects' responses to the 139 statements in the 'life style' section of the questionnaire. The remaining ten variables are individual items selected from this section of the questionnaire. All ten of these variables are related to subjects' attitudes and behavior toward shopping, advertising and the mass media but fail to load on any of the twelve factors in the original principal components factor analysis.

Additional information regarding the attitude of the consumer activist toward shopping and the mass media was obtained through examination of responses to statements which failed to load on any of the twelve factors. In the shopping area, consumer activists tend to agree that the durability of clothing is more important than style and indicate that they prefer to shop alone. The CAC member is also less bothered than are MFP subjects when approached by sales clerks in a retail store. They are also less likely to overspend when shopping and spend less time discussing products and brands with friends. The activist also considers other variables more important than price when making a purchase decision. With regard to the mass media, consumer activists tend to watch more educational television than do mail panel members and also tend to agree that too much television is bad for children.

#### Readership Variables

As evidenced in Table 3, quite different patterns of print media usage emerged in comparing responses to these variables. The most striking difference in the newspaper area is the fact that the consumer activist is a very heavy reader of the Toronto Globe & Mail (a conservative, more establishment newspaper) but reads the Toronto Star far less frequently than does an

average consumer. Much information on the personality of the consumer activist is revealed by the fact that she is a heavier reader of national and international news and of the financial and editorial sections of a newspaper. Her cosmopolitan nature and aversion to advertising is reflected in the fact that she is a significantly lighter reader of the local news, women's pages, and newspaper display advertising. In magazines, the consumer activist is much less likely than is her mail panel counterpart to read traditional women's magazines such as Ladies' Home Journal and Better Homes and Gardens, but is much more likely to be a heavy reader of Canadian magazines and of news and travel magazines such as Time, Newsweek and National Geographic.

#### Listenership Variables

The consumer activist also exhibits a significantly different behavior toward the broadcast media. Table 4 points out clearly that the activist is much less exposed to the broadcast media than are average consumers and would appear to seek relaxation and entertainment in pursuits other than radio and television. Consumer activists have, on average, fewer radio and television sets in the home than do average consumers. The sets which they do own see considerably less use. The consumer activist is exposed to less AM and FM radio

TABLE 3

## UNIVARIATE ANALYSES OF 'READERSHIP' VARIABLES

VARIABLE	T-TEST	SIGNIFICANCE LEVEL	CHI-SQUARE	SIGNIFICANCE LEVEL
Read Toronto Sun Weekdays	1.816	>0.05	*n.c.	
Read Globe & Mail weekdays	-10.078	0.001	96.139	0.001
Read Toronto Star weekdays	6.521	0.001	56.475	0.001
Read Toronto Sun Sundays	4.079	0.001	14.383	0.001
Read Globe & Mail Saturdays	-9.084	0.001	77.899	0.001
Read Toronto Star Saturdays	5.816	0.001	46.329	0.001
Read local and metro news	-2.546	0.05	7.719	0.022
Read national and international news	3.247	0.01	12.958	0.002
Read women's pages	-4.255	0.001	21.521	0.001
Read display advertising	-9.410	0.001	62.864	0.001
Read financial pages	4.132	0.001	20.155	0.001
Read editorial pages	7.405	0.001	45.032	0.001
Read local magazines	-0.581	>0.05	n.c.	
Read contemporary women's magazines	-2.674	0.01	**n.c.	
Read traditional women's magazines	-5.562	0.001	18.198	0.001
Read Canadian magazines	5.216	0.001	35.072	0.001
Read news and travel magazines	4.754	0.001	20.350	0.001
Newspaper exposure	-1.796	>0.05	n.c.	
Newspaper readership	-2.198	0.05	10.300	0.006

\* the crosstabulation analysis was not performed and the chi-square statistic not calculated for those variables which failed to produce a t-statistic which was significant at the 0.05 level.

\*\* the crosstabulation analysis was not performed on this variable because of the very small numbers of subjects in both the MFP and CAC groups who indicated that they had ever read either of the contemporary women's magazines.

TABLE 4

## UNIVARIATE ANALYSES OF 'LISTENERSHIP' VARIABLES

VARIABLE	T-TEST	SIGNIFICANCE LEVEL	CHI-SQUARE	SIGNIFICANCE LEVEL
Number of radios in home	3.389	0.001	7.068	0.030
AM radio listening - weekdays	2.516	0.05	6.980	0.031
FM radio listening - weekdays	2.662	0.01	12.180	0.003
AM radio listening - weekends	2.090	0.05	7.252	0.027
FM radio listening - weekends	2.039	0.05	4.705	0.096
Listen to rock music stations	-8.673	0.001	24.302	0.001
Listen to quality music stations	-5.124	0.001	17.635	0.001
Listen to foreign language stations	-1.343	>0.05	*n.c.	
Listen to quality news stations	5.733	0.001	49.970	0.001
Number of television sets in home	8.716	0.001	71.301	0.001
Hours per day TV on - weekdays	11.266	0.001	104.469	0.001
Hours per day TV on - weekends	11.281	0.001	136.732	0.001
Watch TV variety shows	-10.778	0.001	113.636	0.001
Watch TV talk shows	-11.236	0.001	97.147	0.001
Watch TV game shows	-11.734	0.001	97.362	0.001
Watch TV movies	-6.710	0.001	43.459	0.001
Watch TV afternoon dramas	-8.671	0.001	46.515	0.001
Watch TV situation comedies	-6.963	0.001	49.260	0.001
Watch TV sports	-4.177	0.001	37.725	0.001
Watch TV news	-0.302	>0.05	n.c.	
Watch TV public affairs programs	7.173	0.001	49.303	0.001

\*the crosstabulation analysis was not performed and the chi-square statistic not calculated for those variables which failed to produce a t-statistic which was significant at the 0.05 level

than is the MFP member, both during the week and on weekends. When she does listen to radio, she tends to reject stations which feature a predominantly "music" format in favor of stations which stress quality news reporting and public affairs programs.

The consumer activist is exposed to considerably less television programming than are average consumers. For example, on weekdays, 26.1% of MFP subjects, but only 7.1% of CAC members, are exposed to more than six hours of television programming per day. This relatively light exposure to television is reflected in the fact that the consumer activist is significantly less likely than is the average consumer to watch most types of programs. Both groups are, however, equally likely to be exposed to news programs and the consumer activist is more likely to watch public affairs and documentary programming.

#### Attitudes Toward Advertising

In general, Table 5 shows that consumer activists expressed significantly more negative attitudes toward advertising than did average consumers. This aversion to advertising was exemplified both in reaction to the advertisements which were contained in the questionnaire and in response to certain "life style" statements pertaining to advertising. CAC subjects expressed consistent negative reaction to one of the two advertisements in the questionnaire on all three dimensions (perceived believability, perceived value for the money, motivation to act). Similarly, consumer activists voiced significantly stronger agreement with statements that there is too much advertising on television, that advertising makes people buy things that they don't need, and that they are reluctant to react to television advertising by buying the advertised product.

#### Summary of Univariate Analyses

The univariate t-test and crosstabulation analyses reveal some interesting characteristics of the consumer activist. These may be summarized as follows: - she is considerably better educated and is more likely to be employed outside the home in a management or professional position. - she exhibits more negative attitudes toward shopping and is less likely to be a "bargain hunter". - she would appear to be more interested than is the average consumer in the "functional" attributes of a product. - she uses radio and television for "information" rather than "entertainment". - she exhibits a more negative attitude toward the mass media. - she is likely to exhibit a strong negative opinion of advertising in particular and, likely, of business in general. - she appears to be a much more cosmopolitan individual and more interested in the world around her. - she is more independent and subscribes more to a contemporary view of women in society, rejecting the women's pages of newspapers and women's magazines which may be perceived as presenting a traditional picture of the role of women.

Due to editorial restrictions a complete description of the analyses performed is not contained herein. Although, one should take note that limiting oneself to only a series of univariate significance tests can be misleading and could lead to erroneous conclusions (Tatsuoka, 1971). Thus, a multivariate analysis was conducted and is discussed in depth in another paper (Barnes and Bourgeois, 1975). In summary, the multivariate analysis suggests that not only do the two groups differ significantly on overall mean scores, but that they also relate the variables differently and therefore also 'think' differently.

TABLE 5

#### UNIVARIATE ANALYSES OF 'ADVERTISING REACTION' VARIABLES

VARIABLE**	T-TEST	SIGNIFICANCE LEVEL	CHI-SQUARE	SIGNIFICANCE LEVEL
'perceived believability' housecoat ad	1.114	>0.05	*n.c.	
'motivation to act' housecoat ad	1.917	>0.05	n.c.	
'perceived value for the money' housecoat ad	-2.908	0.01	6.836	0.033
'perceived believability' bedspread ad	4.107	0.001	8.499	0.015
'motivation to act' bedspread ad	5.094	0.001	24.583	0.001
'perceived value for the money' bedspread ad	3.192	0.01	7.022	0.031
"My choice of brands for many products is influenced by advertising"	1.396	>0.05	n.c.	
"Advertising by retail stores is usually misleading"	1.301	>0.05	n.c.	
"There is too much advertising on television"	-3.761	0.001	10.054	0.007
"Advertising can't sell me something that I don't want"	1.262	>0.05	n.c.	
"I never buy products which I see advertised on television"	-4.059	0.001	11.913	0.003
"Advertising makes people buy things that they don't need"	-4.475	0.001	23.478	0.001
"Advertising drives up prices"	-1.133	>0.05	n.c.	
"Information which I get from friends about products is more reliable than the information I get from advertising"	-1.798	>0.05	n.c.	

\* the crosstabulation analysis was not performed and the chi-square statistics not calculated for those variables which failed to produce a t-statistic which was significant at the 0.05 level.

\*\*the first six variables represent factors which were obtained through factor analysis of subjects' responses to the advertisements in the third section of the questionnaire. The remaining eight variables are individual items selected from the 'life style' section of the questionnaire. All eight of these items are related to advertising but failed to load on any of the twelve factors in the original principal components factor analysis.

### Discriminant Analysis

A discriminant analysis was carried out and owing to the extensive list of variables, only those significantly different (at  $\alpha \leq 0.05$ ) variables in the t-test analysis were entered into the discriminant analysis. The purpose here was to identify a subset of these variables which best "discriminates" between the members of the consumers' association and average consumers. Table 6 presents those statistically significant variables

in the discriminant space was relevant to group differentiation. Thus, it was discovered that not only do we have a statistically significant difference between the two groups, but that this difference was also large.

A classification matrix showing the number in each subsample classified into MFP and CAC groups by the discriminant function is also presented in Table 6. The hit ratio or percentage of respondents correctly classified (87.5%) was approximately 22% better than what

TABLE 6  
SUMMARY OF DISCRIMINANT ANALYSIS RESULTS

Variables	Canonical Coefficient	F		Disc. Function	
		Level	Sig.	MFP	CAC
- respondent's occupation	2.458	281.07	0.001	24.402	19.030
- husband's occupation	0.544	8.70	0.01	20.897	19.699
- husband's education	-0.296	4.40	0.05	11.281	11.927
- watch TV public affairs programming	0.275	13.03	0.001	3.162	2.561
- watch TV situation comedies	-0.268	14.71	0.001	2.572	3.158
- type of home	-0.219	7.22	0.01	2.985	3.454
- read newspaper advertising	-0.189	8.92	0.01	4.670	5.083
- watch TV game shows	-0.172	5.04	0.05	4.423	4.799
- exposure to television (weekdays)	0.123	40.77	0.001	5.783	5.514
- readership of Saturday Toronto Star	0.103	4.72	0.05	2.495	2.259
- importance of durability vs. fashion in clothing purchases	-0.106	5.55	0.05	3.140	3.372
- listenership of FM radio (weekdays)	0.079	3.97	0.05	1.007	0.835
- perceived value for the money (housecoat advertisement)	-0.028	7.36	0.01	0.540	0.602
			Constant:	-95.299	-91.111
			Mean Discr. Score:	0.485	-1.700

$$F_{\alpha < 0.001} = 38.40$$

$$\chi^2_{13} = 366.5, \alpha < 0.001$$

$$\omega^2 = 45.1\%$$

### CLASSIFICATION MATRIX

Group	Classified into Group		Percent Correctly Classified
	MFP1	CAC1	
MFP1	420	60	87.5
CAC1	17	120	87.6
Group	Classified into Group		Percent Correctly Classified
	MFP2	CAC2	
MFP2	415	73	85.0
CAC2	26	104	80.0

### EXPECTED AND ACTUAL HIT RATIOS

	Hit Ratio		Differences	Sig. Level
	Expected	Actual		
Classification Sample	65.5%	87.5%	22%	$\alpha < 0.001$
Validation Sample	67.0%	84.0%	17%	$\alpha < 0.001$

The reader should realize that, in general, high statistical "significance" does not necessarily imply a large "magnitude" of difference, especially when the sample size is large. Thus, the significant F-statistics should not be equated with the magnitude of the difference between the two groups. Tatsuoka (1970) provided a measure ( $\omega^2$ ) which measures the magnitude of this difference. Therefore, to measure the extent of differentiation or total discriminatory power,  $\omega^2$  was used. It was found that approximately 45.1% of the variance

would be expected by the proportional chance criterion (Morrison, 1969). However, it should be noticed that a hit ratio calculated in this fashion may suffer from an upward bias (Morrison, 1969). Thus, in order to overcome this shortcoming, the discriminant function derived from the first half of the sample (MFP1 and CAC1) was used to classify the validation sample or the second half of the sample (MFP2 and CAC2). Although this showed a slight upward bias as hypothesized (84.0% vs. 87.5%) the difference was small and the percent of respondents cor-

rectly classified in the validation sample was also significantly better ( $\alpha \ll 0.001$ ) than that obtainable by chance.

Since we have significantly different sample sizes (about one CAC member for every three MFP members), we should also briefly examine the percent correctly classified in each group. Relying only on the aggregate hit ratio could lead to a misinterpretation of the data. Upon examining the hit ratio for each group we find that the results are quite consistent across all groups (87.5%, 87.6%, 85.0%, 80.0%). Thus, the results were not biased towards any one group, as indicated by the consistent hit ratios across all groups. It is, therefore, concluded, from the statistical significance of the results and from the discriminatory power of the function, that the two groups are indeed quite different.

#### Conclusions And Implications

The most evident conclusion from the analysis described above is that the consumer activist, as represented by the CAC member, is, indeed, very different from the average consumer. These results tend to support earlier findings in this area, but the addition of life style and media-related variables permits a much more comprehensive profile of the activist. The description of the consumer activist which has been presented in this paper suggests a number of important implications.

#### Implications for Marketing Management

The percentage of the MFP sample which was misclassified as activists in a discriminant analysis may be considered an estimate of the percentage of the total population who exhibit characteristics similar to those of the CAC member. This analysis indicated that approximately 15% of the population at large are similar to the activist on those variables which were included in the discriminant analysis. This result suggests that the consumer activist may comprise an important and viable market segment for certain marketing purposes. The fact that such a large percentage of the population likely share the views of the activist toward business and advertising would also likely be of some concern to management.

There is also some suggestion that this 15% of the population which may be considered activists may be understated. It is conceivable that the members of the CAC who would take the time to complete and return a questionnaire are the most active of the activists. If this study has, in fact, attracted responses from the most active of activists, then the proportion of the population which shares the characteristics of the consumer activist is likely much larger than 15%.

In addition to constituting a large segment of the market in their own right, there is evidence that consumer activists are also important as a result of their influence on the behavior of others. The fact that the consumer activist exhibits many of the traits which have, in past studies, been attributed to the opinion leader, would also suggest that she is likely an important channel of communication from the marketer to the population at large. A number of studies have reported on an obvious cosmopolitanism demonstrated by opinion leaders (Rogers & Shoemaker, 1971). This characteristic of an interest and orientation outside her immediate community is definitely demonstrated by the consumer activist in her media usage patterns. This media behavior, being oriented more toward information than to entertainment, would also support the finding of Tigert and Arnold (1971) that the opinion leader is an information seeker.

The opinion leader has also been pictured as more self-confident and independent than the average consumer. Such traits are clearly demonstrated by the consumer activist in that she tends not to rely upon advice from others when shopping and prefers to shop alone. The orientation of the consumer activist outside the home, as suggested by more employment and less dependence on radio and television for entertainment, indicates greater opportunity for social participation, another characteristic of the opinion leader (Rogers & Shoemaker, 1971).

Although few studies have concluded that demographic differences exist between opinion leaders and their followers, a number have suggested that the opinion leader is likely to be better educated and to have higher educational and social status (Robertson, 1971; Rogers & Shoemaker, 1971). Again, these characteristics are clearly demonstrated by the consumer activist.

In addition, many of the characteristics of the opinion leader are also exhibited by the innovator and are similarly reflected by the consumer activist (Rogers & Shoemaker, 1971). Engel, Kollat & Blackwell (1973) report that the innovator is likely to exhibit cosmopolitanism, to be extremely well-informed, and not to be especially interested in low prices per se. Again, these are characteristics demonstrated in the profile of the consumer activist developed in this paper.

Such evidence suggests that the consumer activist should be of interest to the marketer not only because she may constitute at least 15% of the population, but also because she is likely to be particularly influential in informing other consumers.

The question will inevitably arise: "how can the professional communicator reach the consumer activist?" What types of information does she seek and where does she look for it? In her purchase behavior, the consumer activist would appear to look for more functional information than does the average consumer. She is less interested in bargains per se, is not likely to rely upon prices as an indicator of quality, and is more likely to consider the functional rather than the psychosocial aspects of a purchase. Her obvious aversion to media advertising indicates that she does not look to advertising for product information. She also does not seek shopping information from friends but, rather, prefers to shop alone. Her self-confidence and independence would suggest that she relies heavily upon her own ability to determine what is the best purchase decision. Her membership in the CAC would also suggest that she makes use of impartial sources of information such as the Canadian Consumer and Consumer Reports to assist her in making purchase decisions.

The number of consumers who share the interests and characteristics of the activist and the potential for influence of the consumer activist make consumerism a force in society that marketers cannot afford to ignore.

#### Implications for Public Policy

The results of this study provide the public policy maker with information on the characteristics of the leaders of the consumer movement in Canada. These findings, however, raise important questions regarding the extent to which the consumer activist is actually representative of the population at large. The difference between the activist and the average consumer which have been presented above indicate that the activist may not be sufficiently similar to consumers in general to represent their interests before policy-making bodies. The issue of whether consumer advocates and activists actually represent the interests of the mass of

consumers has been raised before (Winter, 1972; Barnes, 1975) and the data presented here suggest further that those charged with the responsibility of formulating regulations in the interest of consumers should at least be aware that the groups and individuals who may be lobbying for such regulations may not be totally representative of the broader consumer interest.

While the question of the representativeness of the consumer activist may be raised, her ability to influence other consumers on various issues should not be ignored. The consumer activist may often appear to be representing the interests of a small group, but her opinion leader characteristics suggest that she is also likely to be capable of mustering public opinion concerning those issues in which she is interested. The vocal consumer activist and the associations to which she belongs have proven themselves effective in the past at generating interest in consumer affairs, often through the use of a willing public press. Legislators have also shown themselves generally willing to support proposals for legislation which are made by activists in the interest of consumers in general. A major danger, however, is that legislators and public policy makers will react to activist pressures by passing laws and regulations which may be neither appropriate nor enforceable.

#### References

- W. T. Anderson, and W. H. Cunningham. "The Socially Conscious Consumer," Journal of Marketing, 36 (July 1972), 23-31.
- H. C. Barksdale, and W. R. Darden. "Consumer Attitudes Toward Marketing and Consumerism," Journal of Marketing, 36(October 1972), 28-35.
- J. G. Barnes. "Consumer Protection Programs as Functional Aids to Risk Handling," Working paper 75-6, Queen's University, 1975.
- J. G. Barnes and J. C. Bourgeois. "Profiling the Consumer Activist: Implications for Marketing and Public Policy," Working paper 75-7, Queen's University, 1975.
- J. F. Engel, David T. Kollat and Roger D. Blackwell. Consumer Behavior. 2nd ed. New York: Holt, Rinehart & Winston, 1973.
- T. P. Hustad and E. A. Pessemier. "Will the Real Consumer Activist Please Stand Up," Journal of Marketing Research, 10(August 1973), 319-324.
- T. C. Kinnear and J. R. Taylor. "The Effects of Ecological Concern on Brand Perceptions," Journal of Marketing Research, 10(May 1973), 191-197.
- T. C. Kinnear, J. R. Taylor and S. A. Ahmed. "Ecologically Concerned Consumers: Who Are They?" Journal of Marketing, 38(April 1974), 20-24
- Donald G. Morrison. "On the Interpretation of Discriminant Analysis," Journal of Marketing Research, 6(May 1969), 156.
- Thomas S. Robertson. Innovative Behavior and Communication. New York: Holt, Rinehart & Winston, 1971.
- E. M. Rogers and F. F. Shoemaker. Communication of Innovations: A Cross Cultural Approach. New York: The Free Press, 1971.
- The Roper Organization. "The Cost of Cleaner Air: How Much Do People Want It?" The Roper Report (October 1971).
- M. M. Tatsuoka. "Discriminant Analysis, the Study of Group Differences," Selected Topics in Advanced Statistics, An Elementary Approach, #6, Champaign, Ill.: Institute for Personality and Ability Testing, 1970.
- M. M. Tatsuoka. Multivariate Analysis - Techniques for Educational and Psychological Research. New York: John Wiley & Sons, 1971.
- Douglas J. Tigert and Stephen J. Arnold. "Profiling Self-Designated Opinion Leaders and Self-Designated Innovators Through Life Style Research," Working paper, University of Toronto, 1971.
- D. J. Tigert, J. G. Barnes and J. C. Bourgeois. "Research on Research: Mail Panel versus Telephone Survey in Retail Image Analysis," The Canadian Marketer, 7 (Winter 1975).
- R. K. Winter. "The Consumer Advocate versus the Consumer," Washington: American Enterprise Institute for Public Policy Research, 1972.

## STOCHASTIC WEIGHTS IN MULTIATTRIBUTE DECISION MAKING

Dov Pekelman, University of Pennsylvania  
Subrata Sen, University of Rochester

### Abstract

Most multiattribute models assume that the consumer's brand perceptions, ideal point, and attribute weights do not change from one buying situation to another. However, since consumers express preferences for different brands in different purchasing situations, one or more of these parameters must be changing over time. To explain such multibrand preference, we present a model which allows one of the model's parameters, the consumer's attribute weights, to vary across purchasing situations.

An important problem in marketing is the design of new products or the modification of existing products. Products can be characterized by a set of descriptive attributes such as crispiness, sweetness, and calorie content for dry cereals. Product design, therefore, involves the determination of attribute levels which will result in a successful product. In order to do this, it is first necessary to understand how consumers combine and weight the various attributes in order to arrive at overall preference judgments for the existing brands of a product class.

Traditional economic demand theory is not equipped to aid the marketing researcher in this task because it does not utilize any information on product attributes. It starts with information on product preferences (expressed in terms of an indifference map) which already incorporates all information about product attributes. All that demand theory can say about product attributes is that the preference map would be different if the products were different. Since there is no theory about how product attributes affect preferences, traditional demand theory cannot predict how demand would be affected by changes in one or more attributes of an existing brand or how a new brand would affect the preference structure for existing brands (see Lancaster, 1971, pp. 1-12, for a more detailed discussion of these issues).

In recent years, however, various models of the process by which consumers combine product attributes to arrive at overall preference judgments have been presented in the marketing and decision making literature. The major models are the linear compensatory model (Bass, Pessemier, and Lehmann, 1972, Bass and Talarzyk, 1972, Lehmann, 1971, and Pekelman and Sen, 1974), the conjunctive and disjunctive models (Einhorn, 1970, and Wright, 1973), and various versions of the lexicographic model (MacCrimmon, 1968, Russ, 1971, and Tversky, 1972). Most of the models cited above deal with three parameters of the choice process: (1) the ideal point, defined as the consumer's vector of most-preferred levels for each attribute, (2) the attribute weights, defined as the consumer's vector of saliences indicating the relative importance of each attribute, and (3) the consumer's perceptions of the locations (or attribute levels) of each brand. Most models have implicitly assumed that these parameters are unique for each consumer and do not vary from situation to situation. But is this a reasonable assumption? We attempt to answer this question by exploring the value of a model which allows these parameters to vary across purchase situations.

We start by examining why one might wish to view these

parameters as being stochastic. First, company advertising is often designed to alter the consumer's ideal point, attribute weights, or brand perceptions so as to increase the probability of purchasing the company's brand (see Boyd, Ray, and Strong, 1972). But even in the absence of such advertising, it might be reasonable to allow at least the consumer's ideal point and weight vector to vary across purchase situations. If the consumer's ideal point and weight vector always remained unchanged we would expect the consumer to provide the same preference ranking of the brands at all times. However, when confronted with different purchase situations, consumers typically provide preference rankings which change from situation to situation. For example, Belk (1974) reports that he obtained different likelihoods of purchase for snack items such as potato chips, popcorn, cookies, assorted nuts, and crackers when he varied the purchasing situation. The following examples are illustrative of the purchasing situations presented by Belk to his sample of consumers:

1. You are shopping for a snack that you or your family can eat while watching television in the evenings.
2. You are planning a party for a few close friends and are wondering what to have around to snack on.
3. You are going on a long automobile trip and are thinking that you should bring along some snacks to eat on the way.

Other examples of changing consumer preferences when confronted with different purchasing situations are provided in Bass, Pessemier, and Lehmann (1972), pp. 538-541, Day (1972), p. 284, Green and Rao (1972), pp. 152-180, Hansen (1969), p. 442, and Sandell (1968).

The increasing body of evidence regarding the existence of situational behavior has motivated the recent development of stochastic preference theories (Bass, 1974, and Hermiter, 1973). However, these theories do not explain why a consumer's behavior changes across situations. We suggest, in this paper, a possible explanation. Specifically, we postulate that changes in situations alter a consumer's ideal point or weight vector, resulting in changes in his overall brand preference structure. The changed preference structure manifests itself in changed preference rankings or likelihoods of purchase as the situation is varied.

So far, the discussion has centered around consumers' stated preferences. What about their actual purchasing behavior? For inexpensive, frequently purchased products, diary panel records indicate that most consumers do not always purchase the same brand. Some of the brand switching is undoubtedly due to price variations, unavailability (out-of-stock) of the consumer's most-preferred brand, and word-of-mouth recommendations for the brand that the consumer switched to. However, even when these marketing factors do not exhibit much change (as frequently happens for limited time periods), we observe brand switching on the part of consumers. Again, it seems reasonable to postulate that this brand switching is occasioned by changes in either the ideal point or the weight vector. We should point out that an alternative explanation for such brand switching is pro-



vided by Lancaster (1971), pp. 25-49. However, see Pekelman and Sen (forthcoming) and Hendler (1975) for some problems with Lancaster's approach.

We believe that a model with stochastic parameters (i.e., a model whose parameters can vary from situation to situation) can be useful in explaining situational effects in buyer behavior. In other words, multi-brand preference (or purchase) can be explained fairly well by a model with stochastic parameters. Therefore, the principal objective of this paper is to present a model which allows the consumer's weight vector to vary from situation to situation. To simplify the model, we will continue to assume that his ideal point remains unchanged. We use the stochastic weight model to predict the frequency with which consumers purchase each of the brands in a particular product class. To test the predictive power of the model, we compare our predictions with the consumer's actual purchase frequencies as recorded in a diary purchase panel.

In this paper, we limit ourselves to validating the stochastic weight model. But if we can validate this model, we will obtain a clearer understanding of how consumers utilize product attributes to form preference judgments. This will enable us to make better predictions of consumers' purchases which will aid us in our ultimate objective of helping manufacturers to make better decisions regarding the design of new brands and the modification of existing brands.

The next section consists of a description of our model of how consumers combine product attributes to arrive at overall preference judgments. This preference model is then used to formulate the optimization problem which enables a firm to determine the best combination of attributes for its brands. The preference model is initially described with a deterministic ideal point and weight vector. We then allow the weight vector to be stochastic and provide a numerical example to illustrate the nature of our stochastic parameter model. Next, we provide a formal representation of the model. The paper concludes with a description of the data available to test the model.

### Preference Model Structure

Consider  $q$  brands each characterized by a set of  $n$  attributes. For each attribute, a consumer has a specified preferred position denoted as his ideal point. When the consumer evaluates a brand, he estimates the "distance" between the brand's position on each attribute and his ideal point. Once these distances are estimated for each attribute, the consumer computes an "overall distance" by means of a weighted combination of the individual distances. These ideas can be expressed more formally as follows:  $x_{ij}$  = location of consumer  $i$ 's ideal point on attribute  $j$ , where  $i = 1, \dots, m$ ,  $j = 1, \dots, n$  and  $\bar{x}_i = (x_{i1}, x_{i2}, \dots, x_{in})$ .  $y_{kj}$  = location of brand  $k$  on attribute  $j$  where  $k = 1, 2, \dots, q$  and  $\bar{y}_k = (y_{k1}, y_{k2}, \dots, y_{kn})$ .  $d_{ijk}$  = distance of brand  $k$  from consumer  $i$ 's ideal point on attribute  $j$ , i.e.,

$$d_{ijk} = |x_{ij} - y_{kj}| \quad (1)$$

$w_{ij}$  = weight assigned to attribute  $j$  by consumer  $i$  and

$\bar{w}_i = (w_{i1}, w_{i2}, \dots, w_{in})$ .  $D_{ik}$  = overall distance of consumer  $i$  from object  $k$ , i.e.,

$$D_{ik} = \sum_j d_{ijk}^2 w_{ij} \quad (2)$$

In Appendix A we distinguish between two main types of attributes and discuss the economic implications of each attribute type. Briefly, attribute type A is such that a consumer prefers as much (or as little) of it as possible. The following two automobile attributes are examples of type A attributes: miles obtained per gallon of gasoline and repair expenditures per month. If the upper and lower limits of  $z_j$  (the level of attribute  $j$ ) are  $\bar{z}_j$  and  $\underline{z}_j$ , the consumer's ideal point for type A attributes is located either at  $\bar{z}_j$  or at  $\underline{z}_j$ .  $\bar{z}_j$  is possibly  $\infty$  while  $\underline{z}_j$  is usually 0. The second type of attribute (type B) is such that the most preferred level is not at either extreme. Examples of type B attributes are: the amount of sweetness in a cereal or the amount of carbonation in a soft drink. Essentially, a consumer would generally prefer a cereal with a certain amount of sugar and his preference would decrease if the amount of sugar increased or decreased from that level.

We assume that cost differences involved in producing brands with different levels of type A attributes may be considerable, while for type B attributes it is negligible (see Appendix A for a discussion of this assumption). We concentrate on type B attributes in this paper primarily because we can separate the problem of finding the optimal combination of attributes from the problem of setting the optimal price (again, see Appendix A for a more detailed discussion of this point).

The preference model described above can be used to formulate the firm's optimization problem which enables it to determine the best combination of attributes for its brands. Since the primary concern of this paper is the validation of the stochastic parameter model, we relegate the formulation of the firm's optimization problem to Appendix B.

### The Stochastic Attribute Weight Model

The solution of the optimization problem described in Appendix B depends upon the values of the preference model parameters. These parameters have typically been assumed to remain constant. However, as was argued earlier, it seems more reasonable to allow these parameters to vary across purchase situations. Consequently, we present below a model with stochastic parameters. To keep the model simple, we assume that the consumer's ideal point does not change. However, we allow his weight vector to shift from situation to situation. As the weights vary, the rank orderings, in general, will not remain the same since the  $D_{ik}$ 's (defined in equation (2)) will change for each situation. Changes in the rank orderings can result in changes in the most-preferred brand. The proportion of situations in which a particular brand attains the most-preferred position will constitute our prediction of how often the consumer will choose that brand. To validate the model, we need only to compare these predictions with the consumer's actual choice patterns.

In order to make such predictions, the first step is to determine all the situations that can possibly occur. This is an extremely difficult task since each situation may be characterized by a combination of a large number of events, and different situations may be characterized by different sets of events. This fact also suggests that it would, in general, be infeasible to relate each one of the elements characterizing a situation to choice behavior. For example, situations preceding a choice of a cereal brand may be characterized by the composition of the menu, the time of the day, temperature, appetite, etc. Each of these elements has some implicit relationship to choice behavior, but to investigate them all

would be impractical. However, each situation corresponds to the consumer's choice of a vector of weights which, in turn, leads to the rank order specified by the consumer for the situation. Therefore, by delineating all possible rankings we essentially cover all possible situations.

We are now ready to formalize the model which is based on the assumption that the probability of purchasing brand  $k$  given a set of weights is equal to 1 if  $D_k < D_{k'}$  for all  $k \neq k'$ , and equal to zero otherwise. We start by describing a simple numerical example which illustrates the basic ideas of our model.

#### Numerical Example

Let the distances of three brands (A, B, and C) from an individual's ideal point on three attributes (the  $d_{ijk}$  defined in (1)) be as follows:

	Brands		
	A	B	C
Attributes	1	3	2
	2	1	1
	3	2	3

Let  $w_1$ ,  $w_2$ , and  $w_3$  represent the individual's weights on the three attributes and let the sum of the weights equal 1.0 for simplicity. The first step in the analysis is the specification (for every pair of brands) of the constraints which express the fact that one brand of the pair is preferred to the other. Starting with brands A and B, the constraint specifying that A is preferred to B is as follows:

$$(D_A - D_B) \leq 0 \Leftrightarrow (-2w_1 + w_2 + w_3) \leq 0 \Leftrightarrow w_1 \geq 1/3$$

Similarly,  $(D_A - D_C) \leq 0 \Leftrightarrow (-w_1 + w_2) \leq 0 \Leftrightarrow w_2 \leq w_1$

$$(D_B - D_C) \leq 0 \Leftrightarrow (w_1 - w_3) \leq 0 \Leftrightarrow w_2 \leq (1 - 2w_1)$$

These constraints define ranges of weights within which a certain rank order would hold as illustrated in Figure 3. For instance, the area PQRS in Figure 3 represents: (1)  $w_1 \geq 1/3$ , implying A is preferred to B, (2)  $w_2 \leq w_1$ , implying A is preferred to C, and (3)  $w_2 \geq (1 - 2w_1)$ , implying C is preferred to B. This gives

rise to the rank order  $\begin{bmatrix} A \\ C \\ B \end{bmatrix}$  (which indicates that brand A

is preferred to brand C and brand C is preferred to brand B). The basic logic of the analysis is unaffected if, instead of the City Block metric used in the above constraints, we use the Euclidean metric described in equation (2) (see Pekelman and Sen, 1974, for a discussion of the appropriate metric).

The next step of the analysis involves the prediction of the frequency with which each brand is chosen by the consumer. This requires the determination of a distribution for the attribute weights. Initially, we will assume a uniform distribution of weights. Given this assumption, the choice frequency for each brand for a three-attribute problem is given by the ratio of the sum of the areas where the brand is most preferred to the total area. For example, the computation of  $Pr\{A\}$ , the probability of choosing A, involves first the summa-

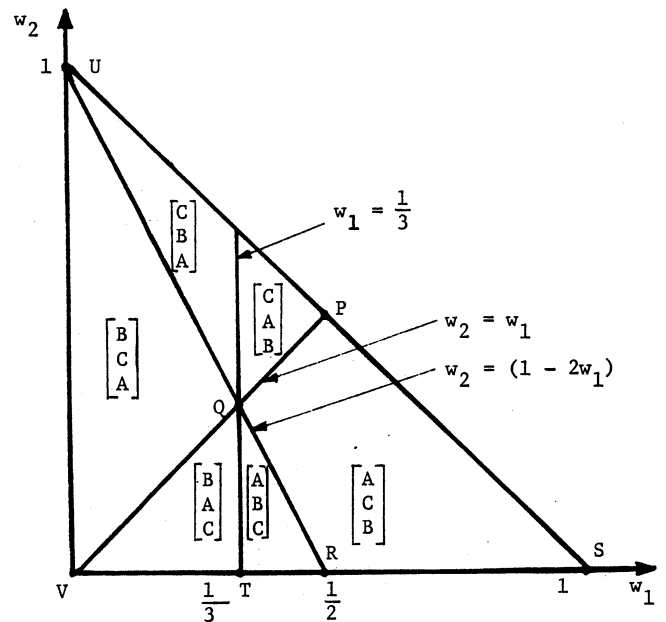
tion of the areas where A is the most-preferred brand: areas PQRS and QTR in Figure 3.  $Pr\{A\}$  is then computed as the ratio of this area sum, PQTS, to the total area, UVS. Using this model, the predicted probabilities are:

$$Pr\{A\} = 0.39, \quad Pr\{B\} = 0.44, \quad \text{and} \quad Pr\{C\} = 0.17.$$

The final step is to compare these predictions with the consumer's actual choice behavior as computed from the record of his purchases in a diary panel. If, for example, the consumer's purchase history indicated that he bought brand A 40 per cent of the time, brand B 45 per cent of the time, and brand C 15 per cent of the time, our predictions ( $Pr\{A\} = 0.39$ ,  $Pr\{B\} = 0.44$ , and  $Pr\{C\} = 0.17$ ) would compare very favorably with actual behavior.

FIGURE 3

RANGES OF ATTRIBUTE WEIGHTS DEFINING PREFERENCE RANK ORDERINGS



We also present a second (and weaker) version of the model in which we attempt to incorporate information about unrealizable rankings since, for a given consumer, not all possible rankings are realizable. For example, if we consider three objects, the consumer may state that two of the six possible rankings have not occurred in the past and are highly unlikely to occur in the future. Such a statement would make sense if, for example, the consumer could not visualize any circumstances in which one of the three objects would be his most-preferred alternative. As shown below, unrealizable rankings modify our brand choice prediction in a simple manner. Clearly, this version of the model amounts to what we may call "partial prediction." Here we predict just part of the choice distribution, the realizable rankings, while the other part of the distribution, the unrealizable rankings, is given as input.

Returning to our numerical example, assume that the consumer states that the rank order  $\begin{bmatrix} B \\ A \\ C \end{bmatrix}$  is unrealizable. This amounts to redefining the "total area" in Figure 3 as (UVS - QVT). The modified predictions are:

$$Pr\{A\} = 0.44, \quad Pr\{B\} = 0.37, \quad \text{and} \quad Pr\{C\} = 0.19.$$

## Formal Representation of the Stochastic Weight Model

The formal representation of the model utilizes the following notation:

$s_u$  = a discrete situation in the situation space  $S$ ,  $u = 1, 2, \dots, h$  where  $h$  denotes the number of situations.

$w_u$  = a vector of attribute weights corresponding to  $s_u$  (one to one).  $w_u$  is an  $n$ -dimensional vector in  $w$ -space where  $n$  equals the number of attributes.

$R_r$  = a particular rank order of all objects,  $r = 1, 2, \dots, q!$  where  $q$  denotes the number of objects.

$W_r$  = the set of vectors,  $w_u$ , leading to the rank order,  $R_r$ . The set  $W_r$  contains  $a_r$  members. In other words,  $a_r$  weight vectors lead to the rank order  $R_r$ . Clearly,  $a_r$  can equal zero.

$c_k$  = event of having object  $k$  ( $k = 1, 2, \dots, q$ ) as the most preferred object.  $c_k$  represents the decision to choose object  $k$ .

$\tilde{R}_k$  = the set of all  $R_r$ 's such that the most preferred brand in  $R_r$  is  $k$ . Let  $b_k$  be the number of members in this set.  $b_k$  will range between 0 and  $(q-1)!$  (In this model, ties are not allowed).

$\tilde{W}_k$  = the set of  $\tilde{R}_k$  defined in terms of attribute weights, i.e.,

$$\tilde{W}_k = \cup_{r \in \tilde{R}_k} W_r \quad (15)$$

$\Pr(s_u)$  = probability that situation  $s_u$  will occur. This is also equal to the probability that the weight vector,  $w_u$ , will occur. Thus,  $\Pr(s_u) = \Pr(w_u)$ .

$\Pr(R_r)$  = probability that rank  $R_r$  would be realized.

$$\Pr(R_r) = \sum_{w_u \in W_r} \Pr(w_u)$$

$\Pr(c_k)$  = probability that decision  $c_k$  would be realized.

$$\Pr(c_k) = \sum_{R_r \in \tilde{R}_k} \Pr(R_r) = \sum_{w_u \in \tilde{W}_k} \Pr(w_u) \quad (16)$$

Our ultimate objective is to compare the computed values of  $\Pr(c_k)$  with the consumer's actual frequency of choosing brand  $k$ .

## Model Testing

The two data sets required for both versions of the model are:

1. Relative frequency of purchase of the various brands.
2. The distances between the ideal points and the locations of the  $q$  brands on each of the  $n$  attributes (the  $d_{ijk}$  of equation (1)).

The first data set can be obtained from existing consumer panel data. The second set of data should be collected from the same consumer group participating in the panel by direct questioning. For the weaker version of our model, there is an additional data requirement, i.e., the specification of the set of unrealizable rankings. This can be obtained by direct questioning in the same questionnaire used to obtain the brand distances.

## The M.R.C.A. Data

The Market Research Corporation of America (M.R.C.A.) has made available to us the coffee purchasing records of a panel of 900 households. The panel was operated by M.R.C.A. between October, 1972 and June, 1973 in Phoenix, Arizona. The panel members were chosen after ensuring that they had purchased both instant and regular coffee. From this panel data, the relative frequency of purchase of the leading coffee brands in Phoenix can be easily computed.

In addition to the panel data, we have available the responses of 542 of the 900 panel members to a series of questions dealing with their perceptions of the 11 leading coffee brands in Phoenix in terms of six product attributes. These 11 brands accounted for 70 per cent of all coffee purchases in Phoenix.

The purchase panel data and the brand perception responses provide us with the information needed to test the stochastic weight model presented in this paper. We cannot test the weaker version of the model because it was not possible to gather data on a complex concept such as "unrealizable rankings" in a mail questionnaire. Such information can probably be gathered only through personal interviews. The questionnaire also provides information on brand preferences. Some limited life-style information is also available. In addition, M.R.C.A. has provided detailed demographic data on each panel family. The data were made available to us very recently. Thus, it has not been possible to present any empirical results in this paper.

## Evaluation of the M.R.C.A. Data

Coffee data are particularly appropriate for testing a model of situational behavior since the choice of instant versus regular coffee is likely to be situation oriented. Further, coffee is a frequently-purchased product which will result in a relatively large number of purchases per family. This will provide us with the sample size necessary to compute market shares for the various brands with a reasonable degree of confidence.

However, it is important to point out some of the potential problems in using panel data to test the stochastic weight model. Our model cannot distinguish between brand switching due to marketing factors (such as price and promotion) and brand switching due to changing attribute weights. Hence, we expect that the model will perform better for those consumers who are not very sensitive to marketing variables and yet switch between brands frequently (see Blattberg and Sen, 1974, for descriptions of such consumers). Even for these consumers, the brand switching could result not only from changing attribute weights but also from changing ideal points, changing brand perceptions due to company advertising, and purchases made by different members of the family (Davis, 1970). For the purpose of testing the stochastic weight model, we will have to assume that these variables do not have a systematic effect on the brand choice frequencies.

Attribute Types And Their Economic Implications

The nature of the two types of attributes is illustrated in Figures 1a and 1b where  $U(z_j)$  equals the utility derived from consuming the amount  $z_j$  of attribute  $j$  (given  $z_i$  for all  $i \neq j$ ). Type A attributes (Figure 1a) are characterized by:

$$\frac{\partial U(z_j)}{\partial z_j} > 0 \quad \text{or} \quad \frac{\partial U(z_j)}{\partial z_j} < 0 \quad (3)$$

for the entire range of  $z_j$ . Type B attributes (Figure 1b), on the other hand, are characterized by:

$$\left. \frac{\partial U(z_j)}{\partial z_j} \right|_{z_j^*} = 0 \quad (4)$$

where  $z_j^*$  equals the ideal level of attribute  $j$ .

There exists another important difference between these two types of attributes. The cost differences involved in producing brands with different levels of type A attributes may be considerable, while for type B attributes it is negligible. To produce brands of cereal with different amounts of sugar within a limited range would usually involve negligible differences in cost. (Naturally, once we move out of this range, we may have technological problems which could increase the cost sharply.) Similarly, the manufacture of soft drinks with different levels of carbonation would not generally involve different costs. However, the addition of safety mechanisms in a car or the improvement of its miles per gallon performance would generally involve a considerable increase in cost.

FIGURE 1a

UTILITY FUNCTION FOR TYPE A ATTRIBUTES

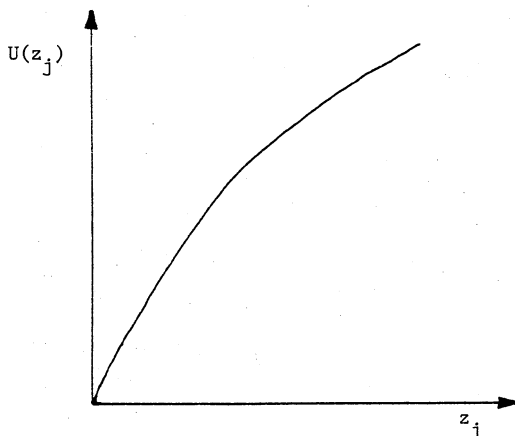
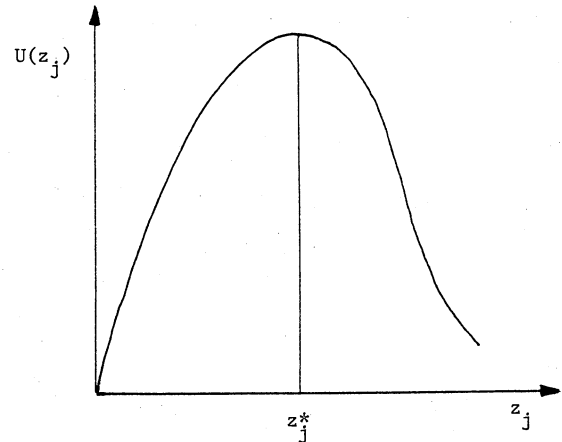


FIGURE 1b

UTILITY FUNCTION FOR TYPE B ATTRIBUTES



These ideas are illustrated in Figures 2a and 2b which show demand and marginal cost (MC) curves for four brands: 1, 2, 3, and 4. The "level of attributes" increases from brand 1 to brand 4. In accordance with our discussion above, we make the following two assumptions: (1) for type A attributes (Figure 2a), each brand has a different MC curve ( $MC_4 > MC_3 > MC_2 > MC_1$ ), and (2) for type B attributes (Figure 2b), each brand has an identical MC curve ( $MC_4 = MC_3 = MC_2 = MC_1 = MC$ ).

For both types of attributes each brand has a different demand curve. For type A attributes, demand increases with the "level of attributes" in the brand, i.e.,  $D_4 > D_3 > D_2 > D_1$ . However, for type B attributes,  $D_2 > D_3 > D_4 > D_1$ . Note that brand 2 has the highest demand while brand 4, with the highest "level of attributes," is superseded in demand by both 2 and 3.

It is obvious, therefore, that if the firm produces a single brand, it would select the one with the highest demand in Figure 2b while it may select any of the levels (depending upon the corresponding MC) in Figure 2a. Essentially, the proposed research examines certain elements of a model for finding the highest demand function when type B attributes are involved.

Given the assumption that there exists an identical MC for all attribute combinations, decisions regarding price and the optimal combination of attributes are separable; i.e., the decisions can be made sequentially. In other words, we wish to identify first the attribute combination which yields the highest demand curve and then determine the price.

FIGURE 2a

DEMAND-COST RELATIONSHIPS FOR TYPE A ATTRIBUTES

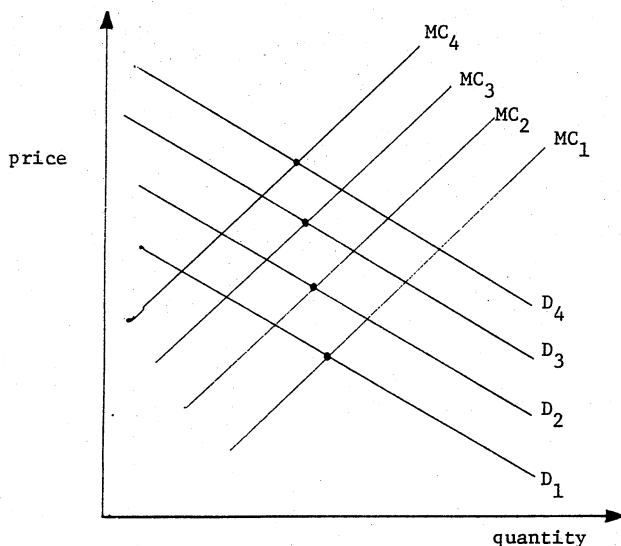
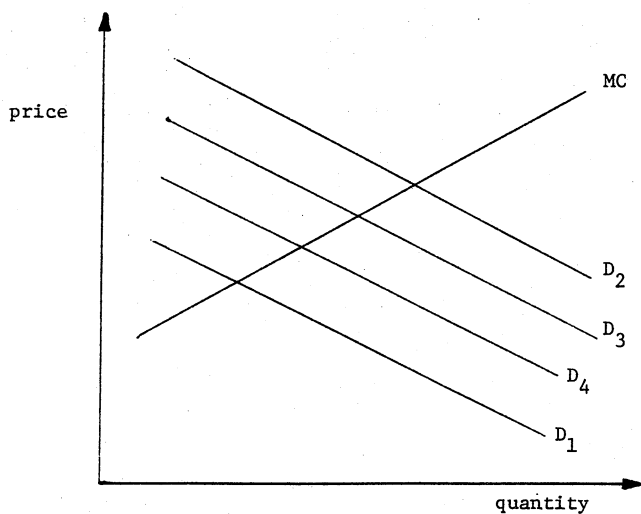


FIGURE 2b

DEMAND-COST RELATIONSHIPS FOR TYPE B ATTRIBUTES



Appendix B

Formulation Of The Firm's Optimization Problem

Let  $P_{ik}$  be the probability that individual  $i$  would purchase brand  $k$  during a specified period of time.

$$P_{ik} = f_{ik}(D_{i1}, D_{i2}, \dots, D_{ik}, \dots, D_{iq}; \quad (5)$$

$$P_1, P_2, \dots, P_k, \dots, P_q)$$

where  $P_1, P_2, \dots, P_q$  are the prices of the  $q$  brands and where

$$\frac{\partial f_{ik}}{\partial D_{ik}, \partial P_{k''}} = 0 \text{ for all } i, k' \text{ and } k'' \quad , \quad (6)$$

$$\frac{\partial f_{ik}}{\partial D_{ik}} < 0 \quad , \quad (7)$$

$$\frac{\partial f_{ik}}{\partial D_{ik'}} > 0 \text{ for all } k' \neq k \quad . \quad (8)$$

Equation (6) assures the separability of the two main problems: the optimal composition of attributes and the optimal pricing policy. And indeed, there is no reason, a priori, to assume that the effect of price (in terms of a change in the probability of purchase) would be different for brands with different overall distances. The function  $f_{ik}(\cdot)$  can therefore be separated as follows:

$$f_{ik}(\cdot) = g_{ik}(D_{i1}, D_{i2}, \dots, D_{iq}) \quad (9)$$

$$+ h_{ik}(P_1, P_2, \dots, P_q) \quad .$$

The function  $g_{ik}(\cdot)$  is discussed in the literature (Schonemann and Wang, 1972 and Stefflre, 1972) where the commonly investigated forms are:

$$g_{ik}(\cdot) = a_i - b_i D_{ik} \quad (10)$$

$$g_{ik}(\cdot) = \frac{a_i}{D_{ik}} \quad (11)$$

$$g_{ik}(\cdot) = a_i e^{-b_i D_{ik}} \quad (12)$$

$$g_{ik}(\cdot) = \begin{cases} 1 & \text{if } D_{ik} < D_{ik'}, \quad k' = 1, \dots, q; \quad k' \neq k \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

In general, the probability of purchase,  $P_{ik}$ , decreases as the distance,  $D_{ik}$ , increases. Given that the consumer distributes his expenditures among brands in a product class according to the specified probabilities, the firm would wish to maximize the expected probability over all consumers. Therefore, if there exist  $r_i$

consumers of type  $i$  and the firm wishes to produce a new product  $\bar{k}$ , then the problem faced by the firm is: Find  $y_{\bar{k}j}$  for all attributes  $j$  such that the expression

$\sum_i P_{i\bar{k}} \cdot r_i$  would be maximized, or, equivalently:

$$\max_i \sum_i r_i f_{i\bar{k}}(D_{i1}, \dots, D_{iq}, D_{i\bar{k}}; P_1, \dots, P_q, P_{\bar{k}}) \quad (14)$$

where

$$D_{ik} = \sum_j (x_{ij} - y_{kj})^2 w_{ij} \text{ for all } k = 1, 2, \dots, q, \bar{k}.$$

Note that the parameters  $r_i$ ,  $y_{kj}$ ,  $x_{ij}$ ,  $w_{ij}$  and, as a result,  $D_{ik}$ , are known. Pekelman and Sen (1974) show how the weights,  $w_{ij}$ , for all  $i, j$  can be numerically derived. The parameters  $y_{kj}$  for all  $k, j$  are generally obtained directly from the various individuals in the sample. The parameters  $x_{ij}$  for all  $i, j$  can be numerically computed using a technique described by Srinivasan and Shocker (1973). The decision variables for the firm are the attribute locations of the new product expressed by  $(y_{k1}, y_{k2}, \dots, y_{kn})$ .

Optimal solutions of the problem posed by expression (14) are difficult to obtain numerically. Consequently, a practical approach is to search for an optimal or near-optimal solution by forming a grid on the attribute space (Shocker and Srinivasan, 1974). Of course, as the number of attributes increases the grid increases exponentially, thereby increasing the search costs considerably. But since the number of attributes is usually no larger than four (Pekelman and Sen, 1974) the search would be feasible. Also note that in general we would not search over the entire attribute space. Usually, technological constraints would restrict the search to a limited region. The problem can be easily extended to deal with the case when the firm wishes to introduce, on the described space, multiple brands. The gain achieved by introducing an additional brand can then be compared to the cost incurred by this action.

#### References

- Frank M. Bass, "The Theory of Stochastic Preference and Brand Switching," Journal of Marketing Research, 11 (February, 1974), 1-20.
- Frank M. Bass, Edgar A. Pessemier and Donald R. Lehmann, "An Experimental Study of Relationships Between Attitudes, Brand Preference, and Choice," Behavioral Science, 17(November, 1972), 532-541.
- Frank M. Bass and W. Wayne Talarzyk, "An Attitude Model for the Study of Brand Preferences," Journal of Marketing Research, 9(February, 1972), 93-96.
- Russell W. Belk, "An Exploratory Assessment of Situational Effects in Buyer Behavior," Journal of Marketing Research, 11(May, 1974), 156-163.
- Robert C. Blattberg and Subrata K. Sen, "Market Segmentation Using Models of Multidimensional Purchasing Behavior," Journal of Marketing, 38(October, 1974), 17-28.
- Harper W. Boyd, Jr., Michael L. Ray and Edward C. Strong, "An Attitudinal Framework for Advertising Strategy," Journal of Marketing, 36(April, 1972), 27-33.
- Harry L. Davis, "Dimensions of Marital Roles in Consumer Decision Making," Journal of Marketing Research, 7 (May, 1970), 168-177.
- George S. Day, "Evaluating Models of Attitude Structure," Journal of Marketing Research, 9(August, 1972), 279-286.
- Hillel J. Einhorn, "The Use of Nonlinear, Noncompensatory Models in Decision Making," Psychological Bulletin, 73(1970), 221-230.
- Paul E. Green and Vithala R. Rao, Applied Multidimensional Scaling (New York: Holt, Rinehart, and Winston, 1972).
- Flemming Hansen, "Consumer Choice Behavior: An Experimental Approach," Journal of Marketing Research, 6(November, 1969), 436-443.
- Reuven Hendler, "Lancaster's New Approach to Consumer Demand and Its Limitations," American Economic Review, 65(March, 1975), 194-199.
- Jerome D. Herniter, "An Entropy Model of Brand Purchase Behavior," Journal of Marketing Research, 10(November, 1973), 361-375.
- Kelvin Lancaster, Consumer Demand (New York: Columbia University Press, 1971).
- Donald R. Lehmann, "Television Show Preferences: Application of a Choice Model," Journal of Marketing Research, 8(February, 1971), 47-55.
- Kenneth R. MacCrimmon, Decision-Making Among Multi-Attribute Alternatives (The Rand Corporation, 1968).
- Dov Pekelman and Subrata K. Sen, "Mathematical Programming Models for the Determination of Attribute Weights," Management Science, 20(April, 1974), 1217-1229.
- Dov Pekelman and Subrata K. Sen, "A Lancastrian Approach to Multiattribute Marketing Models," in Edward M. Mazze, ed. 1975 Combined Proceedings (Chicago: American Marketing Association, forthcoming).
- Frederick A. Russ, "Evaluation Process Models and the Prediction of Preference," in David M. Gardner, ed., Proceedings of the Second Annual Conference (Association for Consumer Research, 1971), 256-261.
- Rolf G. Sandell, "Effects of Attitudinal and Situational Factors in Reported Choice Behavior," Journal of Marketing Research, 5(November, 1968), 405-408.
- Peter H. Schonemann and Ming M. Wang, "An Individual Difference Model for the Multidimensional Analysis of Preference Data," Psychometrika, 37(September, 1972), 275-309.
- Allan D. Shocker and V. Srinivasan, "A Computer-Based Methodology for the Identification of New Product Ideas," Management Science, 20(February, 1974), 921-937.
- V. Srinivasan and Allan D. Shocker, "Linear Programming, Techniques for Multidimensional Analysis of Preferences," Psychometrika, 38(September, 1973), 337-369.
- Volney J. Stefflre, "Some Applications of Multidimensional Scaling to Social Science Problems," in A. Kimball Romney, et al., Multidimensional Scaling, Volume II (New York: Seminar Press, 1972), 211-243.
- Amos Tversky, "Elimination by Aspects: A Theory of Choice," Psychological Review, 79(July, 1972), 281-299.
- Peter L. Wright, "Analyzing Consumer Judgment Strategies: Paradigms, Pressures and Priorities," unpublished paper (College of Commerce and Business Administration, University of Illinois at Urbana-Champaign, 1973).

THE ROLE OF RANDOM WEIGHTS AND RELIABILITY IN THE  
ASSESSMENT OF MULTIATTRIBUTE ATTITUDE MODELS

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

Abstract

In this study multiattribute attitude models using random numbers as weights performed as well as conventional attribute models. Furthermore, there was no relationship between model accuracy and the reliability of attribute beliefs and subject specified importance weights.

Introduction

Since the studies reported by Bass (1972), Sheth (1972), and Talarzyk (1972), research on multiattribute attitude models has increased greatly. A review of some forty-two separate studies by Wilkie and Pessemier (1973) scrutinizes the state of this research and development. A principle issue in the evaluation of this research on multiattribute attitude models has been the inclusion of differential weights (Bass & Wilkie, 1973).

The bulk of empirical evidence indicates that unit-weighted attribute models predict a criterion variable as well as, or slightly better than, subject-weighted models (Wilkie, 1973; Beckwith & Lehmann, 1973; Churchhill, 1972; Moynour & MacLachlan, 1971; Scott & Bennett, 1971; Sheth, 1972). Yet many researchers have argued the necessity of differential weights in consumer attribute models (Wilkie, 1973; Beckwith & Lehmann, 1973; Myers & Gutman, 1974; Scott & Bennett, 1971). The purpose of this study is to analyze the impact of differential weights and measurement reliability in the assessment of multiattribute models of consumer attitudes.

Spectrum of Differential Weights

There exists a natural hierarchy of differential weights for the relationship between a criterion and set of predictor variables. At one end of the spectrum are least-squares derived weights which, when independent and stable, provide the most efficient weights for a set of attributes. Within the context of multiattribute models of consumer attitudes, Beckwith & Lehmann (1973) have demonstrated that least-squares derived weights yield models which predict brand preference better than attribute models using either subject weights or unit weights.

At the other end of the spectrum are differential weights selected from a table of random numbers. This method of weighting attribute models should produce the least effective set of attribute weights for predicting a given criterion. Between these two extremes lie subject-weighted and unit-weighted multiattribute attitude models.

Consumer attitude research is carried out most effectively at the individual level of analysis (Myers & Gutman, 1974). However, this presents a problem since sample sizes needed to achieve stable least-squares weights are difficult to achieve at this level of analysis. To achieve stability approximately 15 to 20 ob-

servations are needed for each attribute included in the model. Since most studies of attribute models involve sample sizes ranging from 5 to 10 observations, least-squares derived models do not offer a viable alternative for either applied usage or benchmarking the relative contribution of models incorporating either subject-stated or unit weights. However, at the opposite end of the spectrum, randomly chosen weights provide an alternative method of assessing model performance, a method that is independent of individual sample sizes. The question becomes: "How much better than models using random weights are the predictions of attribute models using the subject-stated or unit weights?"

Methodology

A sample of 130 adult female shoppers from four local church groups completed a questionnaire concerning shoppers' attitudes. Each group received a self-administered questionnaire with the stated purpose of measuring their "opinions and ideas about various department stores and the shopping process itself." An interviewer was present to answer any questions that the respondents had concerning how to complete the questionnaire.

Respondents' beliefs toward five department stores were measured on a 10-item, 6-interval Stapel scale (Hawkins, et al., 1974). The ten attributes (see Table 1) were chosen from a set used in previous research on department store images using the Stapel scale. The five department stores were selected based on the fact that they were individually well-known in the local community and collectively represented a range of store types. The stores included two high quality regional chains, two medium quality national chains, and one lower quality national chain.

Table 1

STORE ATTRIBUTES USED TO EVALUATE RESPONDENTS' BELIEFS  
TOWARD FIVE DIFFERENT STORES

STORE ATTRIBUTES
Honest
Low Quality
High Priced
Slow Service
Dependable
Helpful Employees
Convenient location for me
Unpleasant
Limited Selection
Friendly

Differential weights were obtained in a two step-process. Each respondent first listed the five attributes she felt were most important in her evaluation of the department stores. Then using a constant sum scale (Hughes 1971) respondents were instructed to allocate 100 points among the five attributes to indicate the relative importance of each attribute.

From each subject's responses, three attribute models were constructed: a subject-weighted model, a unit-weighted model, and a randomly-weighted model using weights drawn from a uniform distribution of random numbers varying from 0 to 100. Each of the multiattribute attitude models is shown below.

$$\text{Subject-Weighted Model: } A_{jk} = \sum_{i=1}^5 I_{ik} B_{ijk}$$

$$\text{Unit-Weighted Model: } A_{jk} = \sum_{i=1}^5 B_{ijk}$$

$$\text{Randomly-Weighted Model: } A_{jk} = \sum_{i=1}^5 R_{ik} B_{ijk}$$

Where:  $i$  = attribute or store characteristic  
 $j$  = store  
 $k$  = female shopper

such that:

$A_{jk}$  = shopper  $k$ 's attitude score for store  $j$

$I_{ik}$  = the importance weight given to attribute  $i$  by shopper  $k$

$B_{ijk}$  = shopper  $k$ 's belief as to the extent to which attribute  $i$  is offered by store  $j$

$R_{ik}$  = the random weight given to attribute  $i$  for shopper  $k$

Two measures were used as criterion variables. One criterion was simply the respondents' rank order of store preferences.

The second criterion was obtained from a constant sum scale by instructing the respondent to distribute 100 points among the five stores to indicate her relative purchases at these stores over the past year. The linear association between the criterion measures and the attitude scores produced by the three attribute models was compiled for each respondent using the Spearman-Rank Order Correlation coefficient.

A second questionnaire was mailed to each participant approximately ten days after the administration of the first questionnaire. The second questionnaire was identical to the first. Test-retest reliability measures of attribute ratings, subject-stated importance weights, rank order of store preferences and subject-stated store behavior were computed for each respondent using the Pearson-Product Moment Correlation coefficient.

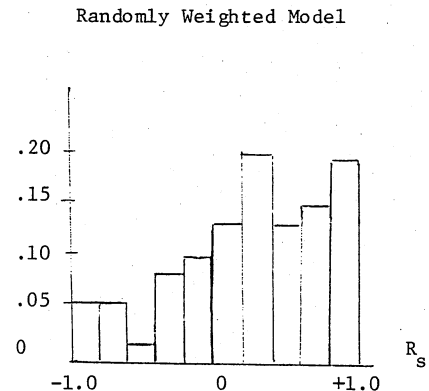
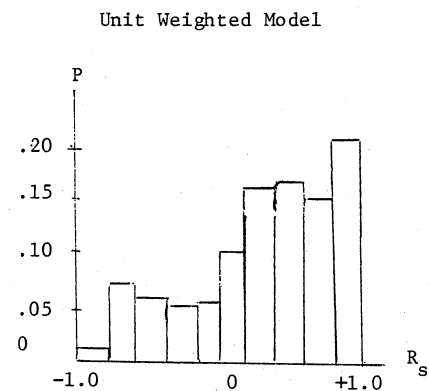
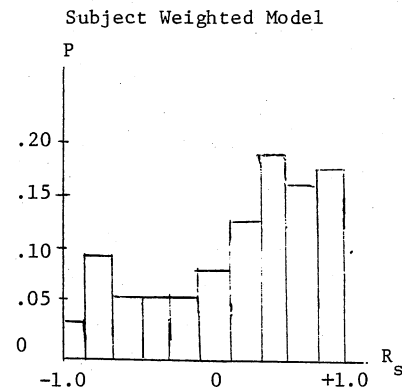
#### Analysis and Results

From the original sample, 114 questionnaires were usable in the analysis, while 70 usable questionnaires were obtained from the second (retest) questionnaire.

The rank order correlation coefficients computed to measure each respondents' linear association between attitude scores and store preferences are plotted for the entire sample in the form of a histogram for each model type specified in this study (see Figure 1). The correlation coefficient distributions produced from subject-stated and unit-weighted attribute models were compared to the distribution produced by the randomly-weighted attribute model using the two-tailed version of the Kolmogorov-Smirnov two sample test. In this case, there were no significant differences ( $\alpha = .10$ )

Figure 1

FREQUENCY DISTRIBUTIONS OF INDIVIDUAL CORRELATIONS FOR ATTITUDE MODELS USED TO PREDICT STORE PREFERENCE



between the distributions generated by the alternative weighting schemes.

The same procedure was followed using subject-stated behavior as the criterion. The distribution of rank order correlation coefficients for each model type is shown in Figure 2. Again, no significant differences were detected at the .10 level when comparing the distributions generated with subject-stated and unit weights with the distribution produced using random numbers as weights. The average correlation between the two criterion variables, store preference and stated shopping behavior was .83 ( $p < .10$ ).

The reliability of the data used to construct these models was assessed with individual measures of test-



retest reliability. The test-retest reliability of belief ratings was computed for each respondent and is displayed in the form of a histogram in Figure 3. The average correlation between belief ratings provided in the test and retest questionnaires was .53 (significant at  $p < .10$ ). A similar distribution is provided for the subject-stated weights in Figure 4. In the case of subject-stated importance weights the average test-retest correlation was .55 which was not significant at  $p = .10$ . The average test-retest correlations for store preferences and stated shopping behavior were .78 and .80 respectively.

Figure 2

FREQUENCY DISTRIBUTION OF INDIVIDUAL CORRELATIONS FOR ATTITUDE MODELS USED TO PREDICT STATED STORE BEHAVIOR

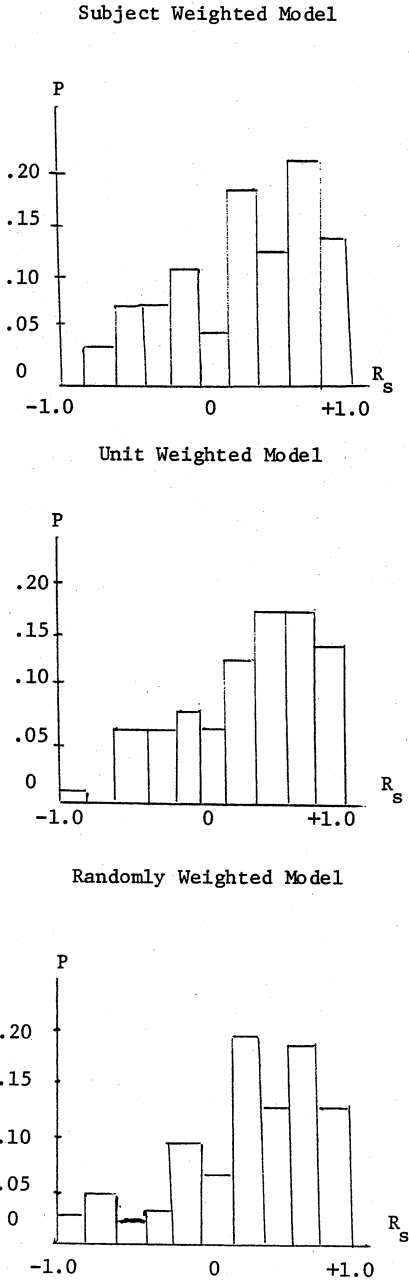


Figure 3

FREQUENCY DISTRIBUTION OF INDIVIDUAL TEST-RETEST CORRELATIONS

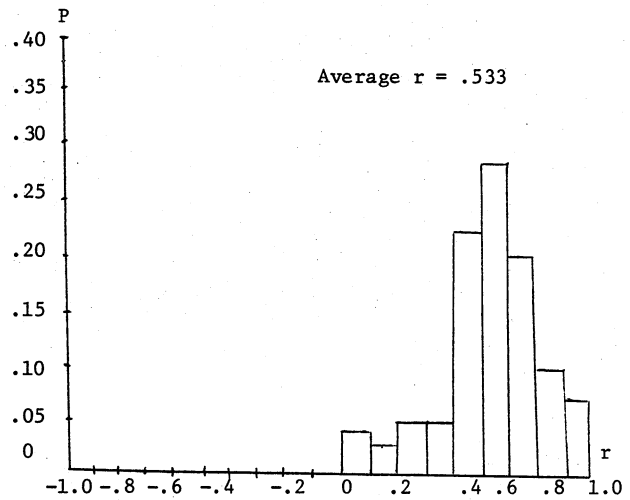


Figure 4

FREQUENCY DISTRIBUTION: INDIVIDUAL RELIABILITY OF STATED IMPORTANCE WEIGHTS

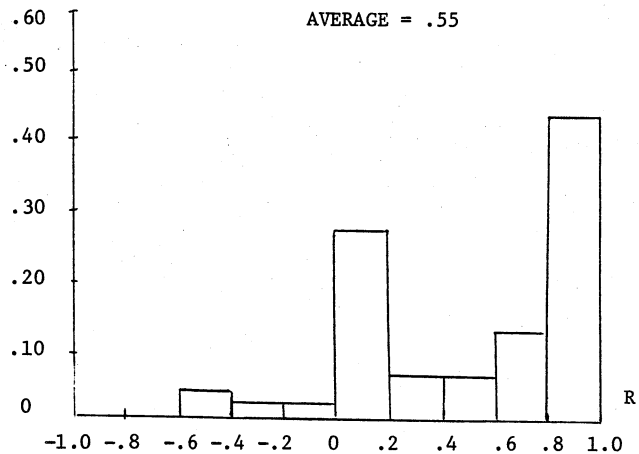


Table 2 was constructed to depict the linear association between model accuracy and test-retest reliability. For each of the model-types shown in Table 2, there was no relationship ( $\alpha = .10$ ) between the model accuracy (correlation between attitude scores and store preferences) and reliability of belief ratings or subject-stated importance weights.

Conclusions

In the aggregate, distributions of rank order correlation coefficients generated by subject-stated and unit-weighted attitude models did not differ significantly from a distribution of correlations derived from models weighted with random numbers. In addition, the reliability of belief ratings and subject estimated importance weights had no impact on the linear association between attitude scores produced by the three model-types and store preferences.

Table 2

CORRELATIONS BETWEEN MODEL ACCURACY AND BELIEF  
AND WEIGHT TEST-RETEST RELIABILITY\*

Model Accuracy	Belief Reliability	Weight Reliability
Subject Weighted	.1062	.0483
Unit Weighted	.0968	.0513
Randomly Weighted	.1378	.1689

\*None of these correlations are significant at  $p=.10$ .

While this study dealt with beliefs toward stores, Holmes (1974) found approximately the same level of test-retest reliability for beliefs related to a set of brands of beer. Thus, while one might expect the reliability of belief ratings to be lower for stores because of departmentized effect on attributes (attributes like service and friendliness could differ greatly between departments within the same store), belief reliability in this study was as high as that reported for a set of products.

The purpose of this study was to evaluate the impact of differential weights and measurement reliability on the predictive accuracy of multiattribute attitude models.

The analysis produced the following results:

1. Randomly weighted multiattribute attitude models perform as well as subject-weighted or unit-weighted models.
2. The accuracy achieved by multiattribute attitude models is independent of the reliability of belief ratings and subject-specified weights which are essential components of the model.
3. A simple rank order measure of store preference is a substantially better predictor of stated shopping behavior than any of the multiattribute models developed in this study.

These findings lead to the general conclusion that multiattribute models are suspect both in terms of their general predictive ability and their ability to aid in our understanding of the variables that influence consumer decisions.

## References

Frank M. Bass, "An Attitude Model for the Study of Brand Preference," Journal of Marketing Research, 9 (February, 1972), 93-6.

\_\_\_\_\_, "Fishbein and Brand Preference: A Reply," Journal of Marketing Research, 9 (November, 1972), 461.

\_\_\_\_\_ and William L. Wilkie, "A Comparative Analysis of Attitudinal Predictions of Brand Preference," Journal of Marketing Research, 10(August, 1973), 262-9.

Neil E. Beckwith and Donald R. Lehmann, "The Importance of Differential Weights in Multiple Attitude Models of Consumer Attitude," Journal of Marketing Research, 10 (May, 1973), 141-5.

Gilbert A. Churchill, "Linear Attitude Models: A Study of Predictive Ability," Journal of Marketing Research, 11(August, 1974), 428-6.

Del I. Hawkins, Gerald Albaum, and Roger Best, "Stapel Scales or Semantic Differential in Marketing Research?" Journal of Marketing Research, 11(August, 1974), 318-22.

David G. Hughes, Attitude Measurement for Marketing Strategies, (Glenview, Illinois: Scott-Foresman, 1971).

Reza Moynpour and Douglas L. MacLachlan, "The Relation Among Attribute and Importance Components of Rosenberg-Fishbein Type Attitude Models: An Empirical Investigation," Proceedings, Fall Conference, American Marketing Association, 1971, 365-75.

James H. Myers and Jonathan Gutman, "Validating Multi-attribute Attitude Models," Proceedings, Fall Conference, American Marketing Association, 1974.

J. C. Nunnally, Jr., Introduction to Psychological Measurement(New York: McGraw-Hill Book Company, 1973).

Jerome E. Scott and Peter D. Bennett, "Cognitive Models of Attitude Structure: Value Importance is Important," Proceedings, Fall Conference, American Marketing Association, 1971, 346-50.

Jagdish G. Sheth, "Perceived Instrumentality and Value Importance as Determinants of Attitudes," Journal of Marketing Research, 9(February, 1972), 6-9.

\_\_\_\_\_, "Reply to Comments on the Nature and Uses of Expectancy-Value Models in Consumer Attitude Research," Journal of Marketing Research, 9(November, 1972), 462-5.

Wayne W. Talarzyk, "A Reply to the Response to Bass, Talarzyk, and Sheth," Journal of Marketing Research, 9(November, 1972), 465-7.

William L. Wilkie and Edgar A. Pessemier, "Issues in Marketing's Use of Multiattribute Attitude Models," Journal of Marketing Research, 10(November, 1973), 428-41.

## OPERATIONALIZING RISK IN MULTIATTRIBUTE DECISION MODELS

Bernard Pras, Ecole Superieure des Sciences  
Economiques et Commerciales

John O. Summers, Indiana University

### ABSTRACT

An operational procedure is developed for treating risk which is compatible with a variety of composition models for multiattribute decisions. Preliminary results utilizing the semi-lexicographic model suggest that risk adjusted brand/attribute ratings may be significantly more predictive of preferences than expected value measures.

### Operationalizing Risk in Multiattribute Decision Models

In spite of the recent interest in composition models for multiattribute decisions (e.g., Russ, 1971; Bettman, Capon and Lutz, 1975; Wright, 1975) and the recognition of risk as an important element in consumer behavior (e.g., Baur, 1961; Cunningham, 1967; Spence, Engel and Blackwell, 1970; Day, 1972) little has been done to incorporate risk in an operational way into any of these models. This is particularly surprising in light of the fact that few would argue that the consumer typically operates under conditions approaching anything like certainty. Where risk has been considered, the tendency has been to attempt to avoid the complications it presents by either discarding subjects with high uncertainty or asking subjects to rate only those brands with which they are familiar. For example, Day (1972) suggests, "... the elimination of the low confidence, instable group, should improve the quality of the relationship." However, Kraft (1972) proposed that a measure of confidence (lack of uncertainty) could be introduced in brand evaluation models at the attribute level.

What is needed is a procedure for operationalizing risk in a manner such that the resulting measures could be easily integrated into any of a variety of composition models for multiattribute decisions. For example, consider the linear additive, conjunctive, and lexicographic models. Underlying each of these classes of models is a different type of evaluation process (Pras and Summers, 1975). Only the linear additive model is compensatory (i.e., poor ratings on one attribute can be "compensated" for by high ratings on another attribute). Furthermore, both the linear additive and lexicographic models utilize measures of the relative "importance" of the determinant attributes while the conjunctive model does not. However, if risk is introduced at the attribute level, the same risk adjusted measures may be utilized for all three types of models. This approach involves the development of risk adjusted brand/attribute ratings which is the direction this research takes. Another potential source of risk, uncertainty concerning the relative "importance" of various attributes, will not be considered here.

The risk adjusted brand/attribute ratings to be developed will consider consumers' risk tolerance (i.e., their willingness to accept risk) for individual attributes and the "acceptability" of their potential values as well as the critical characteristics of the consumer's uncertainty concerning actual brand/attri-

bute levels. The implicit assumption that consumers' risk tolerance varies with attribute "importance" and the "acceptability" of the lowest attribute level considered will be subjected to empirical test. Finally, some preliminary findings will be presented on the pragmatic validity of the proposed risk adjusted brand/attribute ratings.

### The Brand/Attribute Evaluation Process

Composition models for multiattribute decisions implicitly assume consumers evaluate each attribute of a brand separately and then utilize these judgements in some sort of overall comparative brand evaluation process. Consumers generally possess some degree of uncertainty in their judgements and their evaluative perceptions of a particular brand/attribute (e.g., how good or poor Mustang is on gas mileage) might be reasonably represented by a subjective probability distribution over the possible attribute ratings.<sup>1</sup> The question remains as to how consumers might collapse the brand/attribute distributions into single measures for each brand/attribute. Is the mean of the distribution sufficient to summarize the critical information or are higher moments required? If more than one moment is relevant, how should they be combined to produce an overall brand/attribute rating? Finally, should the individual consumer's tolerance for risk be incorporated into these summary measures, and if so, how?

To address the above issues, two sets of potential distributions will be considered (Figure). Case 1 presents two distributions with equal means but unequal dispersions. The consumer who is risk-neutral with respect to this attribute might be expected to be indifferent between the two brand/attribute distributions. Decision theorists have frequently used the expected value as a criterion for deciding between alternative courses of action (e.g., Edwards, 1954). However, a risk-taker (one who has a positive utility for risk) should prefer the one with the larger variance and a risk-avoider the one with the smaller variance. That is, risk-neutral consumers would apply zero weights to the dispersion while risk-takers would assign positive weights and risk-avoiders negative weights. But are the mean and variance sufficient? Case 2 presents two distributions with equal means and variances but different skewnesses. It would seem that consumers who are not risk-neutral should not be indifferent between these two distributions. Risk-avoiders should be more sensitive to the downside portion of the distribution and risk-takers more sensitive to the upper side. Hence it would appear that while the mean may be sufficient to capture the essential characteristics of the distribution for those who are risk-neutral, the dispersion and skewness must be considered for those

<sup>1</sup>The basic assumption here is that all relevant brand/attributes are evaluated on some type of excellent to poor or satisfactory to unsatisfactory scale.

who are risk-takers or risk-avoiders. Based on the preceding discussion, a risk adjusted measure or index of the brand/attribute evaluation should equal the mean when the consumer is risk-neutral and should utilize higher moments to adjust this value upward when the consumer is a risk-taker and downward for risk-avoiders.

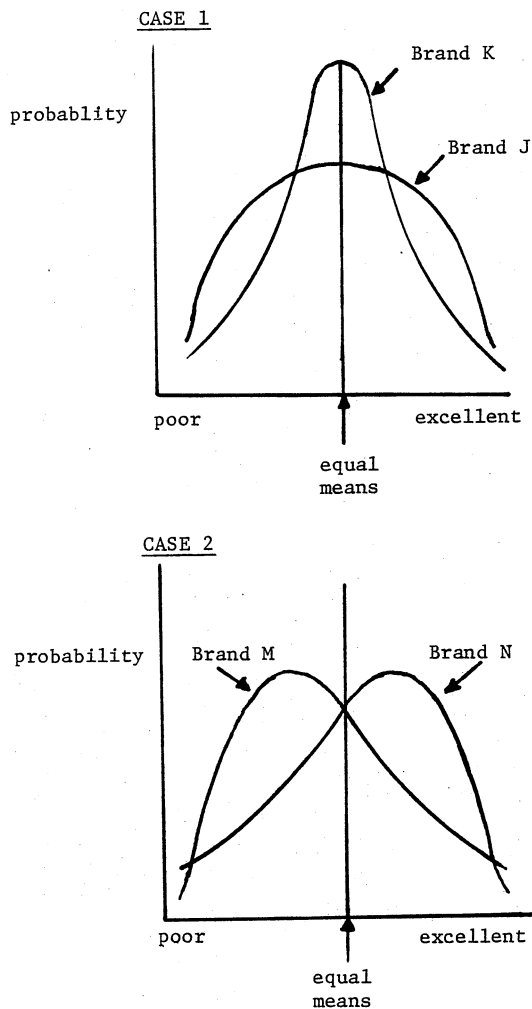


FIGURE 1.  
SOME POTENTIAL BRAND/ATTRIBUTE DISTRIBUTIONS

To this point, risk tolerance has been treated as a personality characteristic rather than a situational variable. While there is a precedent for this in the literature (e.g., Kogan and Wallach, 1964), Slovic (1962) has observed that risk-taking is highly task specific. In the context of attribute evaluation, two factors may affect the individual's risk tolerance: 1) the importance of the attribute of interest, and 2) the "acceptability" of the potential attribute levels. More specifically, it is hypothesized that individuals will have less risk tolerance: 1) for attributes of high importance, and 2) in situations where at least one of the potential attribute levels is "unacceptable."

#### A Risk Adjusted Brand/Attribute Evaluation Measure

Consider the following risk adjusted measure for brand/attribute evaluations:

$$P_{ij} = \mu_{ij} + r_{ik} \sigma_{sij}$$

where  $P_{ij}$  = the risk adjusted index for attribute  $i$  and brand  $j$

$\mu_{ij}$  = the mean of the distribution for brand  $j$  on attribute  $i$

$r_{ik}$  = the consumer's tolerance for risk for attribute  $i$  with respect to the range of possible ratings ( $k$ )

$\sigma_{sij}$  = the semi-standard deviation of the distribution with respect to the mean. This will be the downward semi-standard deviation if the consumer is a risk-avoider and the upward semi-standard deviation if the consumer is a risk-taker.

Had the variance or standard deviation been used instead of the semi-standard deviation, the skewness of the distribution would not have been taken into account. Furthermore, the index has the additional advantage of being expressible in a simple functional form. The assignment of  $r_{ik} = 0$  to those consumers who are risk-neutral with respect to the attribute suggests they only utilize the mean. Assignment of  $r_{ik} = -a$  to a consumer who tolerates no risk, and  $r_{ik} = +a$  for a consumer who is willing to take any risk provides for the maximum adjustments of the mean ratings in the appropriate direction. The value of "a" will be selected to provide maximum goodness of fit to preference data.

#### The Data

To test the above assumptions regarding the sensitivity of risk tolerance and to provide evidence regarding the usefulness of the risk adjusted ratings, data were collected from 40 undergraduate volunteers concerning their evaluations of automobiles. The attributes to be rated were selected from an initial extensive list developed from individual interviews and a literature search. "Determinant" values (Alpert, 1971) and measures of "similarity of meaning" (Pras, 1973) among attributes were obtained from a pre-test group of 20 students. The attributes included in the final list were such that each had a high "determinant" value and there was little overlap of meaning among the attributes. The test subjects rated each of five brands on the final eight attributes (retail price, ease of handling and ride qualities, styling, dealer service, performance, accommodation, durability, and gas consumption). They also evaluated the relative "importance" of the attributes on a constant sum scale and completed risk tolerance items for each attribute. The brands rated were specific to each subject to insure reasonable familiarity with all alternative brands (Pras and Summers, 1975).

#### Measures of the Means and Semi-Standard Deviations

Operationalizing this risk adjusted measure requires that either the subjects directly report their means and semi-standard deviations for all brand/attribute combinations or that they provide sufficient information from which to derive their subjective probability distributions for the same. Since it was anticipated that most subjects would have difficulty comprehending semi-standard deviations let alone estimating them, the latter approach was selected. The specific

approach utilized was a modification of a scale developed by Woodruff (1972). Basically, the procedure involved requiring the subjects to assign points to each possible rating (on a 17-pt. poor to excellent scale) for each brand/attribute combination in such a way as to reflect the relative likelihood that the stated brand actually possessed that rating for the specified attribute. These brand/attribute point distributions were easily adjusted to reflect the underlying subjective probability distributions by dividing by the total number of points assigned. The means and semi-standard deviations were then computed from the resulting probability distributions. It should be noted, however, that this basic procedure is not recommended when the total number of brand/attribute combinations is very large because of the great burden it places on the subject.

#### Operationalizing Risk Tolerance

An individual's tolerance for risk might be considered to be a measure of the degree to which he prefers a given payoff with certainty to a lottery of equivalent expected value. To the extent the individual requires relatively higher payoffs on the part of lotteries to be indifferent with respect to certainty payoffs, he may be considered to have less tolerance for risk or a more negative attitude toward risk. This is consistent with Von Neumann and Morgenstern's (1947) introduction and subsequent treatment of the concept, and it forms the basis of Kogan and Wallach's (1964) operational measure of risk tolerance.

While some researchers (Bem, Kogan, and Wallach, 1965; Nisbitt and Gordon, 1967) have treated risk-taking propensity as a personality characteristic, Slovic (1962) observed that risk-taking behavior is situational. Therefore, it should be measured at the same level of specificity as the variable to which it is compared or related. In the context of this study, it might be expected that risk tolerance would be attribute specific. More specifically, consumers may be less tolerant of risk on those attributes which they consider more important in their choice decision. In addition, consumers might be willing to accept more uncertainty if all possible attribute values (those which have nonzero probabilities) are above some minimal acceptable level. Consideration of the above lead to the measurement of risk tolerance at the attribute level under each of two conditions. In the first condition, all possible attribute levels were "acceptable" while in the second condition the lower level was "unacceptable." This enabled the testing of the assumptions regarding the sensitivity of risk tolerance to attribute importance and the "acceptability" of possible attribute ratings.

The basic methodology utilized to measure risk tolerance was a modification of Kogan and Wallach's (1964) procedure. The method requires the subject to assess the probability he would have to have of winning the higher of two attribute levels (in a lottery) in order to be indifferent between a lottery ticket and an attribute level halfway between with certainty. The approximate interval properties of the attribute rating scale were determined by Myers and Warner's procedure (1968). The risk tolerance measure for the first condition (all possible attribute levels "acceptable" as it appeared for the attribute brakes

was:  
You are going to buy a car and have a choice between three models: model A has Fair brakes, model B Good brakes, model C Excellent brakes. You are given the choice between getting B with certainty (good brakes) and getting a lottery ticket with the

certainty of winning either A (fair brakes) or C (excellent brakes).

Your decision of choosing B or the lottery ticket will depend upon your chances of winning model C (excellent brakes). For which chances of winning model C would you be indifferent between getting B (good brakes) with certainty or getting the lottery ticket. Please check the lowest chances of winning C (excellent brakes) that are acceptable to you in order to be indifferent.

\_\_\_\_\_ 0 chances in 10 of getting model C (excellent brakes)  
\_\_\_\_\_ 1  
\_\_\_\_\_ 2  
\_\_\_\_\_ 3  
\_\_\_\_\_ 4  
\_\_\_\_\_ 5  
\_\_\_\_\_ 6  
\_\_\_\_\_ 7  
\_\_\_\_\_ 8  
\_\_\_\_\_ 9  
\_\_\_\_\_ 10

The risk tolerance measure for the other condition differed only in attribute levels utilized (i.e., poor, fair, and good).

At this point it should be noted that a pretest on twenty undergraduates for three different attributes showed that the rating Poor never passed the minimum acceptable level, and Fair was judged to equal or exceed the minimum acceptable level in 57 cases out of 60. Because of the high agreement, Fair was taken as passing this level for all subjects in the major study.

Since the certainty payoff was halfway between the two lottery payoffs, a subjective probability assessment of .5 would indicate equal expected payoff from the certainty and lottery alternatives. Hence a probability assessment of .5 would indicate the subject was risk neutral in this situation and a risk tolerance of 0 was assigned. A probability assessment of 1 for the better attribute value would indicate the subject had no tolerance for risk ( $r_{ik} = -a$ ), and probability of 0 a subject who was willing to take any risk ( $r_{ik} = +a$ ). From the above considerations and the fact that subjective probabilities are natural ratio scaled variables, the following scale for risk tolerance was adopted.

Lowest chances acceptable (Chances in 10 of getting model C)	Risk tolerance indexes	Attitude
_____ 0	+ 1a	
_____ 1	+.8a	
_____ 2	+.6a	Risk takers
_____ 3	+.4a	
_____ 4	+.2a	
_____ 5	0	Risk neutral
_____ 6	-.2a	
_____ 7	-.4a	
_____ 8	-.6a	Risk avoiders
_____ 9	-.8a	
_____ 10	- 1a	

## An Analysis of Risk Tolerance

Two major hypothesis concerning risk tolerance were tested:

1. The higher the value importance of the determinant attribute, the more the consumer is a risk avoider; that is, the lower is his risk tolerance.
2. Individuals have less tolerance for risk when one of the dubious attribute values is unacceptable than when all possible attribute values (those with non-zero probabilities) are acceptable.

Since it was considered that there was potential for an interaction between attribute "importance" and "acceptability," the data were analyzed as a 2 (levels of "acceptability") times 3 (levels of "importance") factorial design with repeated measures. The three levels of value importance were the two most important attributes, the next two attributes, and the four least important attributes. The risk indices were averaged within each of these levels for each subject to provide the 6 basic observations for each subject.

Since the interaction effects were not significant ( $p > .1$ ), the main effects of the two factors are subject to unambiguous interpretation. The main effects of "acceptability" were highly significant ( $F = 25.7$ ,  $p < .001$ ). Furthermore, the results were in the predicted direction. The subjects were substantially less willing to accept risk (mean difference = .22 based on a scaling of risk tolerance from -1 to +1) when one of the potential attribute values was unacceptable. While the main effects of value importance was not significant ( $p > .1$ ), subjects were somewhat more willing to tolerate risk on the four attributes they considered least important. Perhaps this failure to find significant differences was partially due to the fact that each attribute included had a fairly high level of importance to the subjects.

These results suggest risk tolerance is sensitive to the level of "acceptability" of the worst potential outcome. As such it would seem critical to make sure the risk tolerance measure utilized is consistent with the range of potential payoffs under consideration. However, assuming only attributes with a relatively high level of importance are utilized, it may not be necessary to measure risk tolerances for each individual attribute.

### Some Preliminary Results

The previous sections serve to provide a rationale for the operationalization of the risk adjusted attribute ratings. However, the ultimate usefulness of this procedure can only be determined within the context of the multiattribute alternative evaluation models (e.g., conjunctive, linear additive, and semi-lexicographic) for which it was developed. At this time only results for the semi-lexicographic model utilizing the two most important attributes are available. To provide a basis for comparison three variants of the original risk adjusted measure were used in the analysis. These variants were derived by making different assumptions about risk tolerance.

The original risk adjusted measure and the three variants are as follows:

original operational  
definition

$$P_{ij} = \mu_{ij} + r_{ik} \cdot \sigma_{sij}$$

1st variant

$$P_{ij}^a = \mu_{ij}$$

2nd variant

$$P_{ij}^b = \mu_{ij} - a\sigma_{sij}$$

$\sigma_{sij}$  is the semi-standard deviation of the distribution below the mean only

3rd variant

$$P_{ij}^c = \mu_{ij} + r_{ij} \cdot \sigma_{sij}$$

$$r_{ik} = r_{ik} \text{ for risk avoiders}$$

$$r_{ik} = 0 \text{ for risk takers}$$

The original measure assumes that the subjects take a lack of confidence into account but can also in some cases be willing to take risk. Risk-taking propensity interacts with confidence both ways; the subjects may avoid risk or they may want to take risk on some attributes. The first variant assumes that the subjects make their decision on the basis of the first moment of the probability distribution ( $r_{ij} = 0$  for all attributes); that is, they are not influenced by a lack of confidence or by their propensity to take risk. This is in effect a riskless attribute rating measure. The second variant is based on the assumption that subjects are not willing to tolerate any uncertainty in their ratings and therefore that they are going to downgrade all alternatives. This is equivalent to assuming  $r_{ij} = -a$  for all attributes in the original measure. The third variant assumes that the subjects can only be sensitive to their lack of confidence about low attributes. The  $r_{ij}$  value (risk tolerance), when not negative, is set equal to zero.

The measures used for evaluating these variants were the average Spearman rho correlations between each subject's stated preference ordering of those cars he considered "acceptable" (i.e., those to which he assigned a nonzero purchase probability) and the rank orders predicted from the semi-lexicographic model utilizing each variant in turn. The semi-lexicographic model assumes that the consumer evaluates brand/attributes sequentially starting with the "most important" attribute. The successive attributes are only used when the differences among two or more brands on the "more important" attributes are not "significant" (Pras and Summers, 1975). The results for each subject are displayed in the Table.

TABLE 1  
SPEARMAN RANK ORDER CORRELATIONS BETWEEN ACTUAL  
AND PREDICTED PREFERENCES  
(Acceptable alternatives only)

Subject Number	Original Risk Adjusted Measure	1st variant	2nd variant	3rd variant
1	.65 <sup>a</sup>	.65	.67	.60
2	.36	.45	.43	.40
3	.00	.44	-.55	.44
4	-.20	.40	.40	.40
5	.96	.91	.96	.96
6	.89	.85	.84	.83
7	1.00	.63	1.00	.84
8	1.00	.61	.68	.94
9	.36	.12	.10	.36
10	1.00	.93	.95	1.00
11	1.00	.10	.80	.10
12	1.00	.50	.50	1.00
13	1.00	.95	.92	.95
14	.25	-.50	-.12	-.50
15	.64	.20	.34	.45
16	.79	.90	.91	.83
17	-.50	.50	.50	.50
18	1.00	.86	1.00	1.00
19	.89	.20	.80	.10
20	.95	.95	.65	.80
21	.40	-.27	-.80	-.27
22	-1.00	1.00	1.00	1.00
23	.73	.38	.50	.38
24	.10	-.23	-.44	-.23
25	1.00	-1.00	1.00	1.00
26	.25	1.00	-.62	1.00
27	.50	-.83	-.50	-.75
28	.86	.86	.86	.86
29	.81	.31	.31	.31
30	.40	.80	.80	.25
31	.71	.50	.61	.65
32	1.00	1.00	.62	1.00
33	-.30	-.80	-.80	-.80
34	1.00	.86	1.00	.75
35	.96	.71	.79	.96
36	.91	.45	.87	.88
37	.45	.00	.26	.32
38	1.00	.33	.33	1.00
39	.94	.31	.31	.31
40	1.00	1.00	.40	.25

<sup>a</sup>Read: The correlation between the stated preference rank order and that predicted by the semi-lexicographic model using the original risk adjusted brand/attribute measure was .65 for subject 1.

In comparing the results on a subject by subject basis, it can be seen that the original measure performed better than any of the variants in a higher proportion of cases ( $\frac{27}{35}$ ,  $\frac{27}{35}$ , and  $\frac{23}{30}$  for variants 1, 2, and 3, respectively). <sup>34</sup>These proportions are significantly greater than .5 ( $p < .05$ ). Similarly, the mean correlation (across all subjects) was higher for the original measure than for any of the three variants (.62 vs. .43, .46, and .52). However, only the mean differences between the original measure and the first variant was significant ( $p < .05$ ).

Since the original risk adjusted attribute rating performed substantially better (.62 vs. .43) than the riskless measure (1st variant), risk would appear to be an important variable in preference formation. The second variant, which includes the downward semi-standard deviation and assumes all subjects are maximum risk avoiders, provided only a modest improvement (.46 vs. .43) over the riskless case. Since the 3rd variant, which considers only risk avoidance, was still substantially less than the original risk adjusted measure in its predictiveness (.52 vs. .62), it would appear that risk taking also tends to be an important element in explaining preferences.

#### Discussion

One major issue concerning the incorporation of risk in multiattribute decision models is the question of how risk tolerance should be handled. Is it sufficient to treat risk tolerance as a personality variable? Contrary to expectations, the subjects' tolerance for risk was not significantly affected by the differences in importance they attached to the various attributes. However, only attributes of reasonably high importance were included in the analysis. Further research utilizing attributes of more widely varying importance may show consumers are more tolerant of risk concerning attributes of limited importance. Examination of the effects of allowing one of the two potential attribute values (in the risk tolerance items) to be "unacceptable" (while holding the range of potential outcomes constant) show this factor to be very significant. The subjects were less tolerant of risk when one of the potential outcomes was "unacceptable." These results suggest that while it may not be necessary to measure risk tolerance for each attribute individually (assuming each attribute utilized has a relative high level of importance) it would appear advisable to include both a case where all potential attributes values are "acceptable" and a situation where one potential value is "unacceptable."

While only results for the semi-lexicographic model are presently available, there are indications that the subjects did consider risk in forming their preferences. The original risk adjusted measure provided substantially better predictions of preferences than did the riskless measure (expected values of attribute levels). Although this approach to operationalizing risk in multiattribute decision models shows some promise, much more testing utilizing a variety of models and additional subject populations is necessary to establish its true value. Furthermore, other formulations for risk adjusted brand/attribute evaluations are feasible and should be pursued.

There has been a conspicuous lack of research on procedures for operationalizing risk in multiattribute models for consumer decision making. Hopefully, this study will contribute in a significant way to meeting this need and will serve to simulate further research in the area.

#### REFERENCES

- M. I. Alpert, "Identification of Determinant Attributes: A Comparison of Methods," Journal of Marketing Research, 8(May, 1971), 184-91.
- R. A. Baur, "Consumer Behavior as Risk-Taking," Proceedings, Fall Conference of the American Marketing Association, 1961, 389-98.

D. J. Bem, M. A. Wallach, and N. Kogen, "Group Decision Making Under Risk of Adversive Consequences," Journal of Personality and Social Psychology, 1(May, 1965), 453-60.

James R. Bettman, Noel Capon, and Richard J. Lutz, "Multiattribute Measurement Models and Multiattribute Attitude Theory: A Test of Construct Validity," The Journal of Consumer Research, 1(March, 1975), 31-8.

Scott M. Cunningham, "The Major Dimensions of Perceived Risk," in Donald Cox (ed.), Risk Taking and Information Handling in Consumer Behavior (Boston: Division of Research, Harvard Business School, 1967, 82-108).

George S. Day, "Evaluating Models of Attitude Structure," Journal of Marketing Research, 9(August, 1972), 279-86.

W. Edwards, "The Theory of Decision Making," Psychological Bulletin, 51(1954), 380-417.

N. Kogan and M. A. Wallach, Risk-Taking: A Study in Cognition and Personality, (New York: Holt, Rinehart and Winston, Inc., 1964).

Frederic B. Kraft, "An Analysis of the Brand Evaluation-Brand Choice Relationship," unpublished doctoral dissertation, Indiana University, 1972.

J. H. Myers and G. W. Warner, "Semantic Properties of Selected Evaluation Adjectives," Journal of Marketing Research, 5(November, 1968), 401-11.

R. E. Nisbett and A. Gordon, "Self-Esteem and Susceptibility to Social Influence," Journal of Personality and Social Psychology, 5(March, 1967), 268-76.

Bernard Pras, "Predictive Qualities of Linear and Nonlinear Evaluation Process Models," unpublished doctoral dissertation, Indiana University, 1973.

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

Frederick A. Russ, "Evaluation Process Models and the Prediction of Preference," Proceedings, Second Annual Conference Association for Consumer Research, 1971, 256-61.

P. Slovic, "Convergent Validation of Risk Taking Measures," Journal of Abnormal and Social Psychology, 65(1962), 68-71.

H. E. Spence, J. F. Engel and R. D. Blackwell, "Perceived Risk in Mailorder and Retail Store Buying," Journal of Marketing Research, 7(August, 1970) 364-9.

J. Von Neuman and O. Morgenstern, Theory of Games and Economic Behavior, (Princeton, New Jersey: Princeton University Press, 1947).

R. B. Woodruff, "Measurement of Consumers' Prior Brand Information," Journal of Marketing Research, 9(August, 1972), 258-63.

Peter Wright, "Consumer Choice Strategies: Simplifying vs. Optimizing," Journal of Marketing Research, 12(February, 1975), 60-67.



# PSYCHOGRAPHIC RESEARCH IN A CROSS CULTURAL NONPRODUCT SETTING<sup>1</sup>

James F. Engel  
Graduate School, Wheaton College (Illinois)

## Abstract

A survey was undertaken of members of Christian churches in seven major cities of Brazil. The intent was to measure basic lifestyle for the purpose of uncovering felt needs for spiritual growth. This paper reports on modifications of psychographic methodology necessary for cross cultural application and presents major contrasts between lifestyles within two denominations.

Psychographic profiling has now found its way into the standard research repertoire (Wells, 1975). But nearly all published applications have, first of all, been undertaken in a North American or Western European setting. This raises the important question of the applicability of this type of psychological scaling methodology in developing countries which are characterized by substantial lower levels of literacy and differing ways of responding to questionnaire interrogation (Schramm, 1972). In addition, the majority of reported studies have focused on consumer products or services (Wells, 1975), but there is every reason to believe that psychographic research can contribute much to understanding of behavior in other areas of life.

A major survey was undertaken among urban-dwelling Protestant Church members in Brazil during 1974. One of the primary objectives was to assess basic lifestyle--activities, interests, and opinions--for the purpose of uncovering felt needs for spiritual growth. It became necessary to modify the usual psychographic methodology, however, because of important cultural differences. This paper reports on these methodology considerations and presents the sharply contrasting lifestyle profiles of members of the Assembly of God and Lutheran Churches.

## Methodology

This survey was commissioned by Editoria Betania, a major publisher of Christian books, headquartered in Belo Horizonte, Brazil. While there has been a substantial measure of sales success with a few leading volumes (The Cross and the Switchblade and Run Baby Run are of special significance), market opportunities still are far from fully capitalized upon. Part of the problem is reliance on translation of North American books (Engel, 1974). Management recognized the need to initiate an all new series uniquely adapted to the lifestyle within the Brazilian Church. In addition, it is a well-known fact that there are many segments within the broader church which totals over 12 million (Read and Ineson, 1973). Therefore, greater attention must be paid to market segmentation, especially by denomination.

Because the most rapid church growth is occurring in metropolitan areas, the survey was confined to the leading cities of Brazil. The survey was designed and implemented by the Wheaton Graduate School staff in cooperation with SEPAL (Overseas Crusades), a mission

organization specializing in research and strategy for the Church of Brazil. The guiding theoretical framework was the model of spiritual decision processes developed by Engel and Norton (Engel and Norton, 1975), which, in turn, is an adaptation of the Engel, Kollat, and Blackwell model (Engle, Kollat and Blackwell, 1973).

## Sample

Brazil is one of the few developing countries which has a sophisticated national census. This census also provides unusually thorough data on religious background of the population (Read and Ineson, 1973). From these data it was possible to draw a random sample of churches which closely paralleled the proportions in the major denominations in each of the following cities: Belem, Belo Horizonte, Curitiba, Fortaleza, Porto Alegre, Rio de Janeiro and Sao Paulo.

Within each church, an attempt was made to draw a sample of members at random from church rolls in proportion to the size of the membership. This required the cooperation of pastors who serve as a potent gatekeeper, and cooperation was not 100 percent. The result was some minor deviation of the final denominational breakdowns from the originally projected sample. The sample totalled 2,625, and it closely paralleled the expected attributes in terms of age, education, and other demographic characteristics.

## Questionnaire and Interviewing Methods

An unspoken assumption underlying all forms of survey research used in North America is the availability of a literate populace. Although the majority of urban-dwelling Brazilians are claimed to be literate, such is not the case. The most that can be said is that they are semiliterate and can understand elementary textbook Portuguese. This, of course, makes questionnaire development and administration a much more difficult task than it normally would be. It was decided that an interviewer must be present to assure proper understanding of each question. Where necessary, explanations had to be provided, and, of course, it was nearly impossible to standardize these added explanations so that there are no significant variations in the question stimulus presented from one respondent to the next.

The presence of an interviewer, however, only accentuates a pervasive response bias encountered in most developing countries--a tendency to give the "expected answer." This, of course, is always a problem in survey research, but it reaches such proportions in some parts of the world that it is nearly impossible to design valid surveys other than those calling for strictly factual data. This type of response bias cannot be fully eliminated by questionnaire wording, because it reflects a deep seated "people orientation" where the motivation by the respondent is to help the interviewer by giving the expected response.

It was decided in view of these conflicting aspects of interviewer usage to employ a modified group interview format in which respondents came to each church in groups but took the questionnaire individually and anonymously. Interviewers were volunteer Christian workers and seminary students who read each question one-by-one, ascertained understanding of its content,

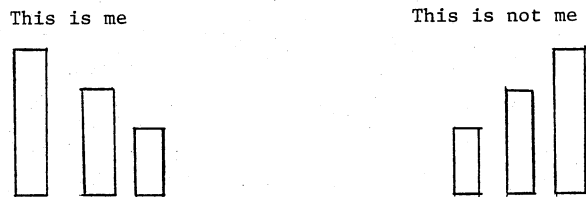
<sup>1</sup>This is an excerpt from The Brazilian Evangelical (Communications Research Reports, Wheaton College Graduate School, April 1975).

and moved to the next question accordingly. The patterns of responses which resulted seemed to verify the success of this method. The highest percentage of non-response did not exceed 10 percent. In addition, there was a good distribution of responses across various categories of agreement which reflected the fact that most apparently avoided giving what seemed to be the expected answer. The guarantee of anonymity, then, seemed to achieve its intended purpose.

Another difficulty is the fact that Brazilians have more of a tendency than North Americans to think in terms of "black and white" rather than in shades of agreement or disagreement. Therefore, a standard five to seven point scale normally utilized in AIO questions was not applicable. To overcome this factor, three different types of lifestyle questions were used.

First, respondents were given a list of various types of behavior such as smoking, drinking, and reading the horoscope. Then they were asked to indicate whether a "good Christian" can or cannot undertake each item of behavior. This inventory was significant in that Christians in the developing countries have a tendency to develop a lifestyle which is legalistic (characterized by "thou shalt not").

A second set of questions was a 28 item AIO inventory. It was developed in part from the standardized inventories used by others (Wells and Tigert, 1971). In addition, a group of questions was designed to focus more directly on issues of concern for this study. To overcome the problem of replying only in "black and white", questions were asked in two phases. First, the respondent was asked to indicate, "is this you" or "is this not you" in response to such items as "my tendency is to worry about the possibility of not having enough money" or "I think that I have more self-confidence than the majority of the people I know". The second phase was, "how much is this you?" and this was gotten by having them fill in a box as follows:



This method proved to be quite workable. A partial validity check provided by similar questions worded in both positive and negative form showed that respondents were not just blindly checking the first or second box and giving the expected answer.

Finally, 15 questions were included which indicated the extent to which help is needed in various areas of life such as learning to study the Bible, handling finances, and so on. Respondents checked whether or not they needed "much help", "some help", or "no help".

One additional validity check was provided in another set of questions which assessed readership of a list of books. One title in the list was a dummy, and it was checked by only one percent of those interviewed. This is further indication that the methods used avoided, at least to a large extent, the tendency to give the expected answer.

#### Tabulation and Analysis

Generally, adequate statistical packages are not readily available in third world countries. Therefore, it was necessary to utilize the facilities of Wheaton College and Illinois Institute of Technology. Our experience

has shown that the best method of analysis is question-by-question cross classification with statistical significance assessed by chi square and the contingency coefficient. This is confirmed by Wells and others (Wells, 1975). Therefore, cross classification was undertaken with each lifestyle question against various demographic breakdowns in the sample. Some of the more useful analyses were differentiation of members of one denomination from another, analysis of differences between readers and non-readers of certain key books, and analysis by age category.

#### Results

Some of the most interesting data was provided by analysis of lifestyle differences between various denominations. Of particular interest were the sharp differences between members of the Assembly of God and the Lutheran churches.

Demographic differences between these two churches were not pronounced, and reliance on demographics alone would have masked essential differences. There were minor differences in age in that the median age in the Assembly was between 25 and 34, whereas the median age of the Lutheran Church was under 25. In addition, median educational attainment of the members of the Assembly of God was primary school not beyond fifth grade, as contrasted with 6-9 grades with its Lutheran counterpart. Other demographic factors did not show statistically significant differences.

Before discussion of the lifestyle data, it should be pointed out that the Lutheran church is predominantly centered in the southern cities of Porto Alegre and Curitiba. Because these cities are distinctly European in outlook and lifestyle, one would expect some real differences between the more traditional lifestyles of the Assembly of God. The Assembly has made its greatest inroads among newer residents migrating from rural areas into Sao Paulo and Rio de Janeiro. The Assembly of God is entirely indigenous in its religious expression, whereas the Lutheran church has until recently at least been dominated by European leadership. As one might expect, the growth rates of the Assembly of God are dramatically higher than the Lutheran church which shows real evidence of stagnation (Read, 1965 and Read and Ineson, 1973).

#### The Inventory of Legalism

Table 1 shows the extent to which the Assemblies and the Lutherans differ on those things that a "good Christian can do."

TABLE 1  
Those Things a "Good Christian Can Do"\*

Behavior	DENOMINATION	
	Assembly of God	Lutheran
Smoke	3.0%	55.9%
Drink	3.5	47.6
Dance	2.9	56.4
Gamble at the state lottery	6.8	60.6
Read the horoscope	11.8	35.2
Play football (soccer)	10.7	91.0
Attend football games	5.9	92.4
Attend football games on Sunday	4.7	83.3
Go to the cinema	4.2	87.8
Go to the cinema on Sunday	3.7	81.5
Watch television	25.2	96.9
Listen to popular music	31.4	92.0
Listen to classical music	50.0	95.2

\* All differences are significant at the 0.05 level.

It is interesting to note that the Bible for the most part does not lay down clear guidelines in any of these areas of behavior with the exception of avoiding of drunkenness and excess. Therefore, these differences are not explainable by reference to the extent to which each church takes the Bible literally.

In reality, the Lutheran church reflects more of an European outlook than Brazilian. The Assembly of God member, on the other hand, is more likely to be a new Christian who has turned sharply from a previous way of life. It is quite to be expected that old ways of behaving would be avoided voluntarily. This also is reinforced strongly by the teaching of the church, whereas this does not appear to be as true among Lutherans.

Whatever the causes, these differences cannot be disregarded. These churches represent sharply different segments, and literature and other types of discipleship materials cannot overlook the almost polar opposite lifestyle profile. The Lutheran would quite likely filter out materials produced for the Assemblies, and vice versa. Given such obvious differences, the sophisticated American marketer no doubt would react in disbelief when it is pointed out that market segmentation has been almost completely disregarded in Christian book publishing, and, sad to say, this is largely the case world wide (Engel, 1974). A book is translated and then distributed to everyone. Needless to say, sales volumes reflect this disregard of marketing commonsense.

The AIO Profile

There were statistically significant differences between these two churches on only 12 of the 28 AIO questions, and data are presented in Table 2.

TABLE 2  
Self-Descriptions of Brazilian Christians on AIO Questions\*

	DENOMINATION	
	Assembly of God	Lutheran
I am not as concerned as I should be with the needs of my non-Evangelical neighbors	52.0%	39.8%
I like to listen to or attend football games	11.9	43.4
I buy things on credit	60.4	47.7
Many times people come to me for counsel about their problems	64.0	54.0
Praying together is part of our family life	75.3	61.6
Often I help the needy	68.5	56.5
I have close non-Christian friends with whom I share my feelings and concerns	42.0	52.8
I do not find it difficult to accept change	48.7	61.0
This month I have tried to lead a nonbeliever to faith in Christ	60.7	22.6
In case of sin in my life, I think I know how to deal with the problem	49.9	60.7
I feel able to speak confidently with a nonbeliever about his need of Christ	76.4	59.5
I have to describe myself as a defeated Christian most of the time	8.1	16.1

\* All differences are significant at the 0.05 level.

Some interesting differences in Christian outlook and behavior are reflected in Table 2. Those Christians

who characterize themselves as evangelical traditionally are more concerned with outreach to the nonbeliever. Clearly this is the case among the Assembly of God. A remarkable number are actively engaged in evangelism, Christian counseling, and social concern ministry. At the same time, a majority feels inadequate in meeting the needs of neighbors, thus indicating a felt need for Christian growth in this area. The Lutheran, on the other hand, is much less engaged in Christian outlook and tends toward a more secular lifestyle as data in both Tables 1 and 2 indicate.

It is interesting that these churches do not differ much on other dimensions such as daily Bible reading, worry about not having enough money, ratings of self confidence, and concern about children and their upbringing.

Felt Needs for Change

Table 3 lists those areas of the Christian life in which respondents indicated they "need much help." In all but

TABLE 3  
Felt Needs for Change

Response	Percentage Needing "Much Help"	
	Assembly of God	Lutheran
Learning to study the Bible	56.3%	21.6%
Sharing my faith	52.8	21.4
Raising my children as Christians	48.7	24.0
Having family devotions	41.6	24.4
Maintaining my Christian testimony in the home	41.8	17.5
Keeping my temper	37.8	16.9
Learning to help others solve their problems	36.7	21.0
Maintaining a Christian testimony on the job	36.0	16.1
Learning to love others	32.6	14.2
Finding my special function in the church	33.5	19.3
Overcoming sexual temptations	37.1	10.0
Making more money	28.6	22.5
Understanding Christians of other groups and churches	28.6	14.2
Handling finances*	16.7	13.9
Birth control	13.7	5.8

\* Not significant at the 0.05 level. All other differences are significant.

one area, the differences between denominations are substantial. It can be said that Lutherans, on the whole, are relatively less concerned with spiritual growth, at least insofar as this inventory is concerned. This is not surprising given a more secular lifestyle. There is a strong indication that religion is seen mostly as just another activity, whereas the member of the Assembly of God embraces his Christianity as a way of life. Concern about Christian growth is central and not peripheral. The implications for Christian book marketing are obvious: the Assembly of God is a fertile market segment whereas the Lutheran Church offers far less potential.

## Discussion

The three sets of psychographic questions revealed sharp differences between the two denominations discussed in this paper. The Lutheran church shows a far more secularistic outlook, and there is evidence that interest in spiritual growth is relatively minimal. On the other hand, Christianity reaches to the very heart of the lifestyle of members of the Assembly of God, and there is both felt need for growth in most areas of life and openness to change. Of the two churches, the latter is a far more fertile market segment. Interestingly, other data showed that over 60 percent of the Assembly of God membership are consistent readers of books, even though educational attainment is minimal. The books utilized, however, are mostly in the "photo-novel" and comic book format. Bookstands are flooded with literature written for the semiliterate. On the other hand, Christian books are largely translated North American titles, which assume at least high school education and a lifestyle very different from the Brazilian. Therefore, the titles now available have relatively little appeal to this important church, and some drastic changes in marketing practices are required. The need now is to produce Brazilian titles, authored by Brazilians, speaking to the felt needs of the Assembly of God members with full cognizance of their distinctly legalistic Christian lifestyle.

On a broader level, this study has demonstrated the utility of psychographic research in a developing country. In conjunction with the Wheaton Graduate School Research Division, it has also been demonstrated that psychographic research, modified for the culture, can be used in such countries as Spain, Jordan, Japan, Ecuador, and Thailand. Obviously appropriate modifications must be made, but the methodology is by no means confined to a North American setting.

## References

- James F. Engel, "Understanding the Market for Christianity--A Report on Asian Christian Strategy," in Scott Ward and Peter Wright (eds.), Advances in Consumer Research, Vol. 1 (Urbana, Illinois: Association for Consumer Research, 1974) 505-15.
- James F. Engel, David T. Kollat, and Roger D. Blackwell, Consumer Behavior, Rev. ed. (New York: Holt, Rinehart & Winston, Inc.) 1974.
- James F. Engel and H. Wilbert Norton, What's Gone Wrong With the Harvest? A Communication Strategy for the Church and World Evangelization (Grand Rapids, Mich.: Zondervan, 1975).
- William R. Read, New Patterns of Church Growth in Brazil (Grand Rapids, Mich.: Wm. B. Eerdmans Publishing Co., 1965).
- William R. Read and Frank A. Ineson, Brazil 1980: The Protestant Handbook, (Monrovia, California: MARC, 1973).
- Wilbur Schramm, Survey Research in Developing Countries (Unpublished manuscript, Stanford University, 1972).
- William D. Wells, "Psychographics: A Critical Review," Journal of Marketing Research, 12(May, 1975), 196-213.
- William D. Wells and Douglas J. Tigert, "Activities, Interests, and Opinions," Journal of Advertising Research, 11(August, 1971), 27-35.

CONSUMER DECISIONS TO REDUCE OR STOP USING PRODUCTS  
AND SERVICES: PRELIMINARY RESULTS OF A NATIONWIDE STUDY

Douglass K. Hawes, University of Wyoming  
Roger D. Blackwell, The Ohio State University  
W. Wayne Talarzyk, The Ohio State University

ABSTRACT

That sphere of consumer behavior involving the decision to cease or reduce consumption of a product or service has been essentially neglected in the published literature. It is apparently assumed that once a consumer becomes brand- or store-loyal, he or she will remain in that category (with some stochastic departures) ad infinitum.

This paper outlines several possible loyalty reduction or cessation situations which have not been addressed in the research literature. The results from part of a nationwide study of 2000 married consumers' participation in, and attitudes toward, leisure time pursuits dealing with this problem are discussed.

In an era of exploding new household formations and unlimited resources to produce new products, consumer research and marketing strategy usually focused on how to attract new customers to new products. As that era ends, consumer research must increasingly focus on the issue of market retention as well as market growth.

Many questions emerge when considering market retention strategies which may be answered by consumer research. What types of products or services are most frequently abandoned by consumers?<sup>1</sup> What are the distinguishing demographic, lifestyle, and attitudinal characteristics of consumers who reduce or cease consumption of a product category or of a brand within a product category? What are the circumstances that lead to the transition from consumption to reduced or ceased consumption? What are the reasons given by consumers for ceasing or reducing consumption of a product category or of a brand within a product category?

This paper reports preliminary results of a study of one type of consumer decision--the decision by consumers to stop or reduce their participation in leisure time pursuits. The study is exploratory, focused upon services rather than tangible products, and confined to one type of decision. It may well provide insights, however, into the general topic of disconsumption--an area of consumer decision making which has received scant attention in the consumer behavior literature.

Specifically, this study seeks to answer the question:

What are the characteristics of persons who reduce or stop consumption of leisure time pursuits, and what are the reasons for the reduction or cessation of consumption?

Background of the Study

There has been little published research on the reasons

<sup>1</sup>The term "product" is used in this paper in the Kotlerian sense of physical object, service, person, place, organization or idea.

why consumers cease to use a brand or product, but there has been considerable research on the topic of brand and store loyalty. Some of this literature is helpful in providing a framework for the analysis of product or brand disloyalty.

An overview of the topic of brand loyalty is found in Consumer Behavior, Second Edition (Engel, Kollat, Blackwell, 1973). A number of journal articles have also surveyed the field (Horton, 1973; Sheth, Park, 1973), particularly the pros and cons of various models of brand choice behavior (Aaker, 1972; Bass, 1974; Blattberg, Sen, 1973; Charlton, French, 1972; Ehrenberg, Goodhardt, 1970; Herniter, 1973; Jones, 1970; Jones, 1973; Kraft, Granbois, Summers, 1973; Montgomery, 1968; Schendel, Summers, Weiss, 1968; Sheth, Park, 1973). Finally, several studies have suggested focusing attention on specific categories of input variables as means to a better understanding of brand loyalty (Carmen, 1970; Chance, French, 1972; Cohen, Houston, 1972; Newman, Werbel, 1973; Weber, Hansen, 1972).

This review of the literature supports the inferences drawn above, that interest in the topic of brand loyalty (of "brand consistent behavior") (Kallick, Nearby, Shaffer, 1973), as a form of repeat purchasing behavior (Jacoby, Kyner, 1975; Tarpey, 1974), has been heavily "front end loaded." That is, emphasis has been primarily on (1) the development of brand loyalty, either single brand or (more recently) multi-brand (Ehrenberg, Goodhardt, 1970; Olson, Jacoby, 1973; Sheth, Park, 1973), (2) those characteristics distinguishing brand- or store-loyal consumers from non-loyal consumers (Carmen, 1970; Lessig, 1973), and (3) distinguishing higher-frequency purchasers of products from lower-frequency purchasers (Barker, Trost, 1973; Wheeler, 1973). Downstream maintenance marketing strategies aimed at retaining currently loyal consumers (Kotler, 1973) in the face of efforts by competitors to cause brand- or product-type switching have not been researched.

While brand loyalty/repeat purchase behavior and its correlates have been studied intensively, published research indicates almost no interest in usage reduction/cessation. In previous research, interest ceased when the consumer crossed that threshold into the "brand-loyal" (or "store-loyal") category for a product-type. An assumption in these models, it would appear, is that once a consumer becomes "brand-loyal," he will remain in that category (with some stochastic departures) ad infinitum.

Little attention has been directed to the following schematized situation:

.....A-A-B-C-A-B-A-A-A-A.....

Initial trial  
of brand A of  
prod. type X

"Loyal" to  
Brand A

A-A-D-A-A-D-A-D-D-D-D-D-D-D.....etc.

Loyalty transition region "Loyal" to Brand D of prod. type X

Examples of this situation would be switching from one brand of pantyhose to another, or, in store-patronage loyalty terms, the downhill skier shifting loyalty from one ski area to another (from Aspen to Vail, for instance).

Similarly, even less attention has been given to the problem of product-type switching, let alone brand switching within the product category. Product-type switching over time might be schematically illustrated as follows:

.....A\_x-A\_x-B\_x-C\_x-A\_x-B\_x-A\_x-A\_x-A\_x.....

Initial trial of Brand A of prod. type X Loyal to Brand A\_x

A\_x-D\_x-A\_x-D\_x-D\_x-A\_x-D\_x-D\_x-D\_x-D\_x-D\_x...

Loyalty transition region Loyal to Brand D of prod. type Y

An example would be the regular coffee drinker who switches, more or less permanently, to tea drinking--perhaps on doctor's orders. Similarly, the blended whiskey (or scotch) drinker who switches to wine would be a case in point, as would be the hobbyist who switches from stamp to coin collecting or the tennis player who switches to golf.

Considerable attention has been given to the heavy-versus-light user problem, and how to (hopefully) move light users into the heavy-user category. Kotler and Levy (1971) have also discussed the situation of "intentionally" moving users back into the light-user category through demarketing efforts. Few studies have addressed the problem of the heretofore "heavy" user who, in spite of efforts to keep him in that category reduces his consumption and falls back into the "light half." Schematically, this might be illustrated as follows:

.....A-B-A-B-B-A-B-A-A-A-A-A-A-A-

Initial trial of Brand A of prod. type X Heavy brand loyal user

-A-A-A-A--A--A--A--A--A--A--A--A--A--A....

Transition region Light brand loyal user

As sketched, this consumer (or group of consumers) remains "brand loyal" while reducing his consumption. Obviously, the same type of problem could be exacerbated by a weakening of brand loyalty or even by a total switch to a different brand--perhaps one which in some ways partially compensates for the reduction in product-type usage frequency. Examples here might be the heavy alcoholic beverage drinker who prefers "getting on the wagon" slowly rather than "cold turkey," or the golfer who finds himself unable to play as often as he once did.

Finally, the last distinct situation is that of the complete cessation of consumption of the product type or generic product category--schematically,

.....A-A-B-C-A-B-A-A-A-A-A-A-A-A-A/

Initial trial of brand A of prod. type X Loyal to Brand A

Cessation of use of prod. type X or generic product category X\_x

The previously avid downhill skier or tennis buff who completely stops engaging in the activity is an example. Similarly, the motorcycle devotee or the power boat enthusiast who stops participating in the activity are cases in point. The cigarette smoker or alcoholic beverage drinker who stops "cold turkey" are familiar examples.

### Research Methodology

The research methodology used in this study involved collecting data on a comprehensive group of leisure-time variables from a representative cross-section of the married United States population. The data encompassed a wide range of leisure-time related variables, in addition to the ones analyzed in this paper.

### Data Collection

The instrument used to collect data was a 20-page questionnaire which requested, in addition to other variables, the following information:

- 1. respondent's participation in leisure-time pursuits in the past and during 1972 (list of 50 pursuits provided);
2. pursuits which had been "completely dropped or stopped" or "reduced" in participation by the respondent during the past 5 years;
3. reasons for stopping or reducing each pursuit listed as "stopped" or "reduced," with up to 5 reasons permitted for each pursuit from a list of 16;
4. demographic variables including respondent's age, education, total household income, occupation, household size and location (rural, suburban, urban).

Other information which was collected included media preferences, extent of agreement with 87 AIO (Activity, Interest, Opinion) statements expected utilization of additional blocks of non-work time, and recreational equipment ownership. Additional papers will report these data when analysis is completed.

A pre-test of the questionnaire was conducted among married graduate students at The Ohio State University to determine and refine the clarity of the questions. A secondary pre-test was conducted using 85 households in the Columbus, Ohio, area. Numerous modifications were made based upon these pre-test results. The final pre-test was administered to 100 households throughout the United States using the sampling procedure described below.

### Sampling Procedure and Response Rate

Two questionnaires were mailed to 1000 households in May, 1973, by Market Facts, Inc. (Chicago, Illinois), from their panel of 45,000 households. Questionnaires for males and for females were sent to households whose names had been recently added to the panel and who had

demographic characteristics proportional to the 1970 Census of the United States.

A total of 603 usable female and 512 usable male questionnaires were returned. The gross response rate was 63 percent while the net usable returns totaled 61 percent.

The demographic characteristics of the respondents were very close to the proportions true of the entire United States on geographic region, urbanization, income, and education, and only slightly higher in age than the adult U.S. population. Families with total household income under \$4,000 were excluded from the original sample, but incomes of respondents were otherwise similar to the income distribution in the total U.S. population.

### Research Findings

The findings concerning the reduction or cessation of participation in leisure-time pursuits are presented in Tables 1 through 4. A brief commentary on these data is presented below, addressed to the following three questions:

1. What leisure-time pursuits were reduced or stopped?
2. What are the characteristics of consumers who reduced or ceased leisure-time pursuits?
3. What are the stated reasons of consumers for reduction of cessation of leisure-time pursuits?

#### Pursuits Stopped or Reduced

People are "disloyal" to activities in which they once engaged. Table 1 shows that 46 percent of males and 54 percent of females either reduced and/or stopped participation in some leisure-time pursuits during the preceding five years.

Table 1

NUMBER OF RESPONDENTS STOPPING, REDUCING, OR MAINTAINING THEIR PARTICIPATION IN LEISURE TIME PURSUITS

Categories	Males (N-512)		Females (N-603)	
	N	%	N	%
1. Only stopped participation in one or more pursuits.	40	8%	38	6%
2. Only reduced participation in one or more pursuits.	94	18%	134	22%
3. Both stopped and reduced participation in different pursuits.	102	20%	154	26%
4. Neither stopped nor reduced participation in any leisure time pursuits.	276	54%	277	46%

The pursuits most frequently stopped or reduced are listed in Table 2. Only those pursuits listed by at least 15 respondents are included. Perhaps different reasons determine stopping than determine reduction in participation. When the frequency of reduction or cessation is recomputed as a percentage of all consumers involved in a leisure-time pursuit (during 1972), it may be seen that the greatest disloyalty among men (20 percent) is for square dancing or other dancing. Among male who are still involved in a pursuit but who have reduced participation, the highest proportion (19.2 percent) is for golf. A hypothesis that might be advanced is that complete cessation may be more related to liking or disliking of an activity, but that reduction for busy people may be more related to time restraints on participation in the activity.

Most of the pursuits reported stopped or reduced in Table 2 are physically strenuous and out-of-the-home. Perhaps these data indicate a trend toward a stay-at-home, sedentary, physically "soft" society. This is consistent with findings that the favorite leisure pursuits of American consumers are listening to music, visiting with friends, reading, and driving around for pleasure or sightseeing (Hawes, 1973).

Table 2

LEISURE TIME PURSUITS MOST FREQUENTLY STOPPED OR REDUCED IN PARTICIPATION

Males-Stopped Participating		N <sup>1</sup>	% <sup>2</sup>
Rank	Pursuit		
1	Playing basketball, football, baseball, softball, volleyball or handball.....	23	9.9%
2	Bowling.....	21	9.3
2	Fishing or Hunting.....	21	7.0
2	Square-dancing or other organized dances.....	21	20.0
3	Camping by tent.....	17	13.0
4	Ice skating, roller skating.....	15	12.0
4	Tennis.....	15	17.6
Males-Reduced Participation			
1	Fishing or Hunting.....	42	12.9
2	Attending movies.....	26	6.2
3	Bowling.....	25	11.1
4	Golf.....	23	19.2
5	Attending sporting events (as a spectator).....	19	5.7
6	Camping by tent.....	16	12.2
7	Playing basketball, football, baseball, softball, volleyball or handball.....	15	6.5
Females-Stopped Participating			
1	Bowling.....	40	15.3
2	Ice skating, roller skating.....	29	16.2
3	Horseback riding.....	26	19.5
3	Playing basketball, football, baseball, softball, volleyball or handball.....	26	13.9
4	Square-dancing or other organized dances.....	23	12.5
5	Camping by tent.....	22	17.2
6	Bicycling.....	19	7.3
7	Tennis.....	17	17.7
8	Playing the piano, organ or other musical instrument for pleasure.....	15	8.6
8	Fishing or Hunting.....	15	6.2

Table 2 (cont.)

Females-Reduced Participation		
	N	%
1	44	8.6
2	30	8.2
3	29	11.1
4	25	4.9
5	23	8.8
6	21	8.7
7	20	11.1
8	19	5.5
8	19	7.3
8	19	3.6
9	18	13.5
10	17	5.6
10	17	17.7
11	16	3.2

Notes: 1. N = number of respondents indicating that they stopped or reduced participation in this pursuit.  
 2. N expressed as a percentage of male or female respondents indicating participation in each of the leisure-time pursuits at least once during 1972.

Demographic Characteristics of Disloyal Consumers

Consumers who are "disloyal" to leisure-time pursuits are described in Table 3. That table shows the most significant demographic characteristics of consumers who have stopped or reduced their participation.

Consumers who have reduced or stopped consumption of leisure-time pursuits are "up-scale" demographically. "Disloyal" male respondents tend to be in the 35-44 age group and not in the over-55 group, and tend to be professionals rather than operatives. Females, who stopped/reduced pursuits, tend to be in the highest income category, have attended or graduated from college, are in the 25-34 age group, and do not live in the urban area of SMSA's.

The age and income distributions of respondents indicates that the busy, middle-aged years of affluent families are prime times for cutting back on activities. As people near the retirement age (55 or over), they appear to be more loyal to old, familiar pursuits.

Reasons for Stopping or Reducing Participation

The most frequently reported reasons for reducing or stopping participation in leisure-time pursuits are shown in Table 4. A pattern emerges which is helpful in understanding the decision to reduce or stop participation.

Table 4 shows that among male consumers the most frequent reason for stopping participation are:

1. Just lost interest (153)<sup>2</sup>
2. Don't have enough time anymore (140)

<sup>2</sup>Number represents the frequency that the reason was selected across all pursuits which were reduced to stopped.

3. Have new, different interests and activities (86)
4. My present family situation makes it difficult to engage in the activity (78)
5. People I used to do the activity with have moved away or lost interest (77)
6. Became too expensive (77)

The reasons for reducing participation which appear in Table 4 are similar:

1. Don't have enough time anymore (205)
2. Just lost interest (143)
3. Have new, different interests and activities (104)
4. My present family situation makes it difficult to engage in the activity (102)
5. Became too expensive (94)
6. People I used to do the activity with have moved away or lost interest (84)

While the rankings of these reasons are similar, the difference is that loss of interest is the most frequent reason for stopping an activity, while the most frequent reason for reducing participation is a perceived time constraint.

Among female respondents, Table 4 reveals that the most frequent reasons for stopping participation are:

1. Don't have enough time anymore (187)
2. My present family situation makes it difficult to engage in the activity (169)
3. Just lost interest (163)
4. Have new, different interests and activities (98)
5. People I used to do the activity with have moved away or lost interest (90)
6. Became too expensive (83)

The reasons given by females for reducing participation also emphasize time constraints:

1. Don't have enough time anymore (232)
2. My present family situation makes it difficult to engage in the activity (229)
3. Just lost interest (149)
4. Have new, different interests and activities (108)
5. Became too expensive (108)
6. People I used to do the activity with have moved away or lost interest (103)

This is rather compelling evidence that consumers' time budgets are more important in the decision to stop or reduce leisure-time pursuits than are their money budgets. This is compatible with the position recently advanced by Voss and Blackwell (1974) that marketing



Table 3

DISTRIBUTION OF RESPONDENTS, BY PARTICIPATION LEVEL CATEGORY,  
ACROSS SELECTED DEMOGRAPHIC VARIABLES

	<u>Males</u>		<u>Females</u>	
	Stopped, Reduced or Both	Neither Stopped nor Reduced	Stopped, Reduced or Both	Neither Stopped nor Reduced
<u>Total Household Income<sup>1</sup></u>				
\$ 4,000 -7,999 (low)	23.8*	32.6*	25.2	31.0
\$ 8,000 -11,999 (medium)	28.1	32.6	26.1*	33.6*
\$12,000 or more (high)	48.1**	34.8**	48.7**	35.4**
	100.0% <sup>6</sup>	100.0%	100.0%	100.0%
<u>Education<sup>2</sup></u>				
Attended High School, but didn't graduate	22.0**	31.5**	15.7**	25.8**
Attended College, but didn't graduate	31.9	33.0	44.3	45.1
College Graduate (or more)	23.3	20.1	27.1*	20.7*
	22.8*	15.4*	12.9*	8.4*
	100.0%	100.0%	100.0%	100.0%
<u>Age<sup>3</sup></u>				
Less than 25	11.4	13.4	22.7	19.1
25-34	31.8	27.2	25.8*	20.2*
35-44	16.1*	9.1*	16.8	13.4
45-54	17.4	18.8	17.5	17.0
55 and over	23.3*	31.5*	17.2**	30.3**
	100.0%	100.0%	100.0%	100.0%
<u>Male's Occupation<sup>4</sup></u>				
Professional	16.9*	9.9*	NS <sup>7</sup>	NS
Manager or Administrator	16.4	14.4	↓	↓
Clerical or Sales	13.1	17.6	↓	↓
Craftsman/Kindred	24.5	24.3	↓	↓
Operative	13.1**	20.7**	↓	↓
Laborer, Farmer, Service Worker	16.0	13.1	↓	↓
	100.0%	100.0%		
<u>Population Density/Degree of Urbanization<sup>5</sup></u>				
Rural--under 2500 pop.	NS	NS	14.7	15.9
Urban--25 to 49,999 pop.	↓	↓	13.8	9.7
SMSA 50,000 and over--Central City			33.4	29.2
SMSA 50,000 and over--Urban			27.6**	37.6**
SMSA 50,000 and over--Rural			10.5	7.6
			100.0%	100.0%

Notes:

- \*Column percentages significantly different at the .05 level (Z test of proportion)  
\*\*Column percentages significantly different at the .01 level (Z test of proportion)
1. Cross-tabulation for males significant at .005 level; for females at .004 level (Chi-square test of association).
  2. Cross-tabulation for males significant at .136 level; for females at .004 level (Chi-square test of association).
  3. Cross-tabulation for males significant at .045 level; for females at .004 level (Chi-square test of association).
  4. Cross-tabulation for males significant at .076 level (Chi-square test).
  5. Cross-tabulation for females significant at .062 level (Chi-square test).
  6. Base N's are given in Table 1.
  7. NS--not significant.

Table 4

LIST OF REASONS FOR STOPPING OR REDUCING PARTICIPATION  
GIVEN TO RESPONDENT AND OVERALL FREQUENCY OF  
SELECTION ACROSS ALL 50 PURSUITS

	Males (N=493) <sup>1</sup>		Females (N=585) <sup>1</sup>	
	Stopped	Reduced	Stopped	Reduced
1. The area or facilities closed down . . . . .	12 <sup>2</sup>	13	18	20
2. Just lost interest . . . . .	153	143	163	149
3. Became too crowded . . . . .	30	34	26	27
4. The weather has been unfavorable . . . . .	14	26	12	20
5. Have new, different interests and activities . . . . .	86	105	98	108
6. Moved away from the area of facilities; it's too far to travel . . . . .	39	46	76	79
7. Took up other activities which prevent me from doing both . . . . .	58	72	62	63
8. Had an accident which prevents me from engaging in the activity . . . . .	5	3	19	17
9. People I used to do the activity with have moved away or lost interest . . . . .	77	84	90	103
10. My general physical condition prevents me from doing the activity now . . . . .	69	56	83	91
11. My present family situation makes it difficult to engage in the activity . . . . .	78	102	169	229
12. Became too expensive . . . . .	77	94	83	108
13. The need for the activity went away . . . . .	26	30	34	36
14. Don't have enough time anymore . . . . .	140	205	187	232
15. Became too tiring or exhausting . . . . .	45	33	56	61
16. Other (please specify) . . . . .	17	23	54	60

Notes:

1. Differs from total sample size due to variable completeness of response to the questionnaire.
2. Each respondent was allowed to select up to 5 reasons for each of up to 5 pursuits in which he/she dropped or reduced their participation.

strategists must analyze buying decisions in terms of the time resources required as well as the money resources. Although the popular press has given much attention to alleged reductions in the work week, it may be that non-work time is so filled with other (non-discretionary) activities that consumers must abandon pursuits or products previously enjoyed because of the competing time pressures.

Losing interest in pursuits once enjoyed is also a common experience. This should indicate to marketing strategists the frailty of consumer loyalty for even favored activities. Strategists must continually consider ways to add new dimensions to an activity or to develop new promotions and excitement if the pursuit is not to fade away from the consumers' preferences because of "lost interest."

Several other insights are revealed in Table 4. Males are less deterred by increased travel necessary to continue participation than are females. Accidents and fatigue resulting from participation are also less of a barrier to males than to females. Females, on the other hand, are slightly less discouraged by crowds and unfavorable weather than are males, possibly because of the general female orientation toward indoor activities.

#### Conclusions

This study indicates that there are substantial numbers of consumers who "drop out" of markets for leisure time activities. The persons who reduce or stop participation in leisure time pursuits are not primarily older persons no longer physically able to participate, nor poor consumers no longer able to afford participation. Rather, the "pursuit disloyal" consumers are middle-aged, educated, moderately affluent persons.

The reasons for stopping or reducing leisure-time pursuits are not primarily money related. They are, however, related to perceived lack of time and lack of interest--both of which factors can be influenced by innovative marketing strategy.

#### Discussion

This study deals specifically with consumer decisions concerning services, i.e., leisure-time pursuits. It is recognized that it is not possible to generalize the findings from this study to the entire field of consumer behavior and brand disloyalty. Several suggestions do emerge, however, that may be worthy of further research and conceptual development in an attempt to understand the phenomenon of ending loyalty to a product-type, brand, or retail source.

First, marketing strategists should collect data on cessation and reduction of usage for product categories and specific brands. These data might be reported in a form similar to "lapse rates" in the insurance industry which indicate the proportion of policies which are dropped for all reasons within a time period. Mathematical models could be developed using lapse rates for product categories to indicate the number of new consumers which must be added to retain present volume or to achieve specified growth rates. It should be feasible to develop additional models which balance target growth rates with optimal levels of resource commitment.

Second, reasons for disconsumption should be determined. This type of research should facilitate improved marketing strategies which again might be compared to the concept of "conservation management" in

the insurance industry. Conservation management refers to marketing programs designed to retain present customers rather than to develop new customers. For example, if other product categories are similar to the leisure-time pursuits investigated in this study, the necessity of continually creating new interest in an existing offering would be critical. Even though the product category is already liked and purchased, the firm may need to adopt a strategy of "innovate or perish" if it is to maintain high acceptance. A measure of such strategies would be a "persistence ratio," which could be compared across market segments, marketing programs, and marketing organizations.

Third, consumption and disconsumption decisions should be analyzed in a framework which includes time budgets as well as money budgets. It is apparent that within the constraint of consumer money budgets, a decision to purchase one product must cause a decision to reduce or terminate purchases of another (perhaps unrelated) product. This is also true within the framework of consumers' time budgets. A product which "costs" time from the time budget often generates a decision to reduce or terminate "purchase" of another product or service which also involves consumption of time. Money budgets have no theoretical upper limits, rise very high for some families, have generally increased during recent decades in the United States, and can be temporarily expanded through the use of credit. Time budgets, however, cannot be expanded beyond 24 hours a day or 365 days a year. Time constraints may therefore provide more reasons for understanding disconsumption decisions of some types than have heretofore been recognized.

Finally, analysis of disconsumption decisions may have important dimensions in a demarketing program either for private firms or for government programs. In an age of scarcity of resources, private business firms may wish to move consumers of a product category into non-consumption mode. The firm usually wishes to undertake such a move, however, with the optimal amount of goodwill toward the company, so that the consumer may later begin repurchasing the product or move to substitutes which minimize the firm's losses. Similarly, the government may be in the position of undertaking demarketing programs for gasoline consumption or energy-intensive forms of leisure activities. In either situation, it becomes essential to understand the consumer processes involved in ceasing or reducing use of a product or service formerly preferred.

#### References

- David A. Aaker. "A Measure of Brand Acceptance," Journal of Marketing Research, 9 (May, 1972), 160-167.
- Stephen M. Barker and John F. Trost. "Cultivate the High Volume Consumer," Harvard Business Review, 51 (March-April, 1973), 118-122.
- Frank M. Bass. "The Theory of Stochastic Preference and Brand Switching," Journal of Marketing Research, 11 (February, 1974), 1-20.
- Stewart W. Bither and Stephen J. Miller. "A Cognitive Theory View of Brand Preference," Proceedings. Fall Conference, American Marketing Association, 1969, 280-286.
- Robert Blattberg and Subrata Sen. "Are Stochastic Brand Choice Models Useful Approximations of Buyer Behavior," Proceedings. Combined Fall and Spring Conferences, American Marketing Association, 1973, 241-245.

- James M. Carmen. "Correlates of Brand Loyalty: Some Positive Results," Journal of Marketing Research, 7 (February, 1970), 67-76.
- William A. Chance and Norman D. French. "An Exploratory Investigation of Brand Switching," Journal of Marketing Research, 9 (May, 1972), 226-229.
- P. Charlton and A. S. C. Ehrenberg. "McConnell's Experimental Brand Choice Data," Journal of Marketing Research, 10 (August, 1972), 302-307.
- Joel B. Cohen and Michael J. Houston. "Cognitive Consequences of Brand Loyalty," Journal of Marketing Research, 9 (February, 1972), 97-99.
- Herbert E. Echelberger and Elwood L. Shafer, Jr. "Snow + (X) = Use of Ski Slopes," Journal of Marketing Research, 7 (August, 1970), 389-92.
- A. S. C. Ehrenberg and G. J. Goodhardt. "A Model of Multi Brand Buying," Journal of Marketing Research, 7 (February, 1970), 77-84.
- James F. Engel, David T. Kollat, and Roger D. Blackwell. Consumer Behavior, Second Edition. New York: Holt, Rinehart, and Winston, Inc., 1973.
- Douglass K. Hawes, W. Wayne Talarzyk, and Roger Blackwell. "Consumer Satisfaction from Leisure Time Pursuits," Proceedings. Advances in Consumer Research, Vol. 2, Association for Consumer Research, 1974.
- Jerome D. Herniter. "An Entropy Model of Brand Purchase Behavior," Journal of Marketing Research, 10 (November, 1973), 361-375.
- Raymond L. Horton. "On the Appropriateness of Brand Choice and Brand Loyalty and Dependent Variables in Consumer Personality Studies," Proceedings. Southern Marketing Association, 1973, 147-151.
- Jacob Jacoby and David B. Kyner. "Brand Loyalty vs. Repeat Purchasing Behavior," Journal of Marketing Research, 10 (February, 1975), 1-9.
- Kenneth L. Jensen. "Purchase Frequency and Consumer Perceptions of Marketing Variables," Proceedings. Southern Marketing Association, 1973, 86-91.
- J. Morgan Jones. "A Dual Effects Model of Brand Choice," Journal of Marketing Research, 7 (November, 1970), 458-465.
- J. Morgan Jones. "A Comparison of Three Models of Brand Choice," Journal of Marketing Research, 7 (November, 1970), 466-473.
- J. Morgan Jones. "Testing Probability of Purchase Distribution and Ad Hoc Market Segmentation Schemes," Proceedings. Combined Fall and Spring Conference, American Marketing Association, 1973, 246-249.
- Maureen Kallick, Joseph Nearby, and Jay Shaffer. "The Dimensions of Brand Consistent Behavior," Proceedings. Advances in Consumer Research, Vol. 1, Association for Consumer Research, 1973, 460-462.
- Philip Kotler and Sidney J. Levy. "Demarketing, Yes, Demarketing," Harvard Business Review, 49 (November-December, 1971), 74-80.
- Philip Kotler. "The Major Tasks of Marketing Management," Journal of Marketing, 37 (October, 1973), 42-49.
- Frederic B. Kraft, Donald H. Granbois, and John O. Summers. "Brand Evaluation and Brand Choice: A Longitudinal Study," Journal of Marketing Research, 10 (August, 1973), 235-241.
- V. Parker Lessig. "Consumer Store Images and Store Loyalties," Journal of Marketing, 37 (October, 1973), 72-74.
- Jerry C. Olson and Jacob Jacoby. "Measuring Multi Brand Loyalty," Proceedings. Advances in Consumer Research, Vol. 1, Association for Consumer Research, 1973, 447-448.
- Joseph W. Newman and Richard A. Werbel. "Multivariate Analysis of Brand Loyalty for Major Household Appliances," Journal of Marketing Research, 10 (November, 1973), 404-409.
- David B. Montgomery. "Stochastic Consumer Models: Some Comparative Results," Proceedings. Fall Conference, American Marketing Association, 1968, 409-415.
- Morris Rosenberg. The Logic of Survey Analysis. New York: Basic Books, 1968.
- Dan E. Schendel, John O. Summers and Doyle L. Weiss. "Simulation and Model Testing," Proceedings. Fall Conference, American Marketing Association, 1968, 416-429.
- Jagdish N. Sheth and C. Whan Park. "A Theory of Multi-dimensional Brand Loyalty," Proceedings. Advances in Consumer Research, Vol. 1, Association for Consumer Research, 1973, 449-459.
- Lawrence X. Tarpey, Sr. "A Brand Loyalty Concept--A Comment," Journal of Marketing Research, 11 (May, 1974), 214-217.
- Justin L. Voss and Roger D. Blackwell. "Markets for Leisure Time," Proceedings. Advances in Consumer Research, Vol. 2, Association for Consumer Research, 1974.
- Jean F. Weber and Richard W. Hansen. "The Majority Effect and Brand Choice," Journal of Marketing Research, 9 (August, 1972), 320-323.
- David R. Wheeler. "Brand Switching Behavior of 'Heavy' and 'Light' Product Users," Proceedings. Southern Marketing Association, 1973, 135-140.

AN INVESTIGATION OF THE INCLUSION OF THE EXPLICIT  
IDEAL POINT IN THE MULTI-ATTRIBUTE ATTITUDE MODEL<sup>1</sup>

Kenneth E. Miller, University of Utah<sup>2</sup>

Abstract

Empirical investigations of alternative forms of multi-attribute attitude models have consistently found predictions to be higher with exclusion of the measured ideal point (a useful concept in brand positioning). The ability of the multi-attribute attitude model in predicting rank order preference was significantly improved when individually measured ideal point ratings were included selectively on the basis of whether or not the attributes were directional in nature.

Introduction

Consumer attitude structure models offer a diagnostic tool for measuring brand image, and marketers employing these models can ascertain the position of their brand relative to competing brands. In addition, these models can furnish information concerning the ideal brand location in the cognitive product space as perceived by each consumer. The ideal brand location is that point in the product space which the consumer prefers over all others. With this information, marketers can ascertain the distance their brand is from the ideal brand and take steps to align their product more closely with the perceived ideal. In this way they can position their brand advantageously relative to competing brands.

The multi-attribute attitude model has been widely used by marketers. Wilkie and Pessemier (1973) have discussed it in the following general form:

$$A_j = \left[ \sum_{i=1}^n W_i |B_{ij} - I_i|^K \right]^{1/K} \quad (1)$$

where:  $A_j$  = attitude toward brand  $j$ .  
 $W_i$  = importance of attribute  $i$ .  
 $B_{ij}$  = perceived amount of attribute  $i$  contained by brand  $j$  (beliefs).  
 $I_i$  = ideal amount of attribute  $i$  (measured ideal point).  
 $n$  = number of attributes.  
 $K$  = parameter of the weighted Minkowski space (determines distance measure to be used in calculating the attitude score. Where  $K=1$ , city block distance is used, and where  $K=2$ , Euclidean distance measure is used).

However, there is a lack of agreement as to the form of the multi-attribute attitude model which best predicts consumer preference.

This paper examines two issues related to the use of the model. First is the question of whether the explicit ideal point should be included in the multi-attribute model. Previous research by Ginter (1972) and Winter (1972) found that the model performed

significantly better in predicting preference when the measured ideal point was not included. The question of whether the measured ideal point should be selectively included in the model has not previously been reported.

The second issue was concerned with the performance of the conjunctive model in predicting rank order preference at the individual level of analysis in an optimization context. The conjunctive model simulates choice strategy by assuming that decision makers reject a brand if its perceived attribute performance is below a minimum cut-off on any attribute. Wright (1975) found that decision makers perceived a conjunctive model as a more likely optimizer than several linear compensatory attitude models.

The Data

The data were collected from a mail panel residing in Columbus, Ohio. The initial panel sample of 744 was randomly generated from a list of names in the Columbus telephone directory.<sup>3</sup> Over a period of three months, each respondent completed five questionnaires concerning eight Columbus fast food restaurant chains. During this period an extensive advertising and couponing campaign was undertaken by one of these fast food restaurants.

Responses were gathered from the mail panel on belief and importance components of the multi-attribute attitude model. Scale values ranged from one (very important or high) to six (very unimportant or low). This information was collected for the following seven attributes:

- Taste of the food.
- Speed of service.
- Popularity with children.
- Price.
- Variety of menu.
- Cleanliness.
- Convenience.

Rank order preference was also obtained at each of the five waves of the questionnaire for each individual.

Forms of the Model

Inclusion of the Ideal Point

There is considerable disagreement on the concepts, measurements, and analysis in marketing's use of the multi-attribute attitude model (Wilkie and Pessemier (1973)). One of the consistent empirical research findings is that the predictive performance of the model is significantly better when the measured ideal point is not included (Lehmann (1971), Ginter (1974), Winter (1972)). Exclusion of the ideal point results in the rather dubious assumption that more of an attribute is better than a lesser amount of the attribute: i.e.,

<sup>1</sup>The research reported in this paper was supported in part by the Division of Research, College of Administrative Science, The Ohio State University.

<sup>2</sup>Assistant Professor of Marketing.

<sup>3</sup>The final panel size was 446 as a result of experimental mortality. The demographic or attitudinal characteristics were not significantly different between the final panel and the drop-outs.

where  $A_j = \sum_{i=1}^n W_i B_{ij}$ , then, holding  $W$  constant, the

higher the value of  $B$  the higher (or more favorable) will be the attitude toward brand  $j$ . This model appears to be lacking where attributes are non-directional in nature: e.g., variety of menu, price. Consumers do not desire maximum variety of menu or minimum price (as illustrated below). Attributes such as cleanliness and speed of service are directional in nature. Where attributes are non-directional (more of the attribute is not preferred to less), it is probable that the ideal point is not at the end point of the belief scale and exclusion of the ideal point from the model should result in poorer predictive performance of the multi-attribute model.

The measured ideal point model (IM) incorporates the ideal point as rated by the respondent. The distance metric used in this comparison of alternate forms of the multi-attribute model is  $K=1$ , as use of the city block distance has been found to lead to significantly better predictions of consumer preference (Wilkie and Pessemier (1973)). Model IM has the form:

$$A_j = \left[ \sum_{i=1}^n W_i |B_{ij} - IM_i| \right] \quad (2)$$

where:  $IM_i$  = measured ideal point for attribute  $i$ .

The mean response to ideal brand ratings for each attribute at time period 1 are shown in Table 1.

TABLE 1

IDEAL BRAND RATINGS AT TIME PERIOD 1

Attribute	Mean Response
Cleanliness	1.24 <sup>a</sup>
Taste of the food	1.30
Service	1.45
Popularity with children	1.69
Convenience	1.72
Variety of menu	2.02
Price	2.37 <sup>b</sup>

<sup>a</sup>1=very high (end points of scale altered for attributes)

<sup>6</sup>=very low

<sup>b</sup>1=very low

From Table 1 it is seen that the mean response to each attribute is greater than one (where one is the assumed ideal point in the

$$A_j = \sum_{i=1}^n W_i B_{ij} \text{ model}).$$

The selected ideal point model (IS) incorporates measured ratings of the ideal point only where the attribute appears non-directional. From Table 1 the attributes variety of menu and price are non-directional in nature and the measured ideal points for these attributes, as given by each respondent, were used in Model IS. During the study, a promotional campaign was conducted by one of the fast food brands studied in order to change the popularity of the restaurant. The measured ideal point of popularity with children was also included in the model. For attributes taste of food, cleanliness, service, and convenience, the measured ideal point was not used. Respondents experience difficulty with the concept of an ideal brand (Ginter (1972)) which may lead to the lack of success of models which include the ideal point. However, selective inclusion of the ideal point

combined with concurrent measurement of the ideal point and beliefs toward each brand was hypothesized to improve the predictive performance of the model.

For comparative purposes a model which did not incorporate the ideal point was tested. The model used was that found superior<sup>4</sup> by Ginter (1974). This model (IA) has the form:

$$A_j = \left[ \sum_{i=1}^n W_i |B_{ij} - IA| \right] \quad (3)$$

where:  $IA$  = assumed ideal point (end point on the belief scale, i.e., 1).

Use of this model yields identical rank order attitude scores to the following model:

$$A_j = \left[ \sum_{i=1}^n W_i B_{ij} \right] \quad (4)$$

#### Conjunctive Model

The conjunctive model assumes that decision makers are satisfiers rather than maximizers. Wright (1975) found that decision makers perceive the conjunctive model as a more likely optimizer than several linear compensating attitude models. According to this approach, consumers will choose the brand which meets a satisfactory standard on all attributes. To obtain a rank ordering of brands under this approach, the satisfactory standard for any individual is increased so that brands are evaluated in terms of their minimum performance on each of the attributes. The formulation of this model (used by Pras (1973)) is:

#### Model CONJ

$$A_j = n |B_{ij} - I_i|_{\max} \quad (5)$$

where:  $|B_{ij} - I_i|_{\max}$  = maximum distance between perceived amount of each attribute and the measured ideal amount of each attribute.  
 $n$  = number of attributes.

#### The Hypotheses

- $H_1$ : Model IM does not outperform model IA in the prediction of rank order preference.
- $H_2$ : Model IS does not outperform models IM and IA.
- $H_3$ : Model CONJ does not outperform models IM, IS and IA.

#### The Analysis

For each respondent, at each of 4<sup>5</sup> points in time and for each model, an attitude score for each brand was computed. The attitude scores obtained using each model were rank ordered and then correlated with the rank order preference measure. The mean Spearman rank correlation coefficients across respondents are shown in Table 2.

The Wilcoxon matched signs test was used to test for significant differences in the magnitude of correlations

<sup>4</sup>Ginter (1974) compared alternative models to determine whether ideal points and attribute importance weights should be included.

<sup>5</sup>Rank order preference data was not collected at time period 1.

TABLE 2

## ATTITUDE PREFERENCE CORRELATION

		IA	IS	IM	CONJ	n
Time Point	2	.727 <sup>a</sup>	.745	.728	.598	473
Time Point	3	.744	.766	.763	.588	394
Time Point	4	.751	.764	.742	.642	392
Time Point	5	.755	.775	.753	.619	382

<sup>a</sup>Mean Spearman rank correlation coefficient.

produced by each pair of models. A matrix of significance levels of differences is outlined in Table 3.

Model IM outperformed model IA in two of the four comparisons of the models ( $p < .001$  and  $< .08$ ) with no difference found in the other comparisons. This result is inconsistent with that of Ginter (1974). The attributes used by Ginter, with a household cleaning product, were stain removing power, whitening power, sudsiness, mildness to clothes, mildness to skin, and pollution control. It appears that these attributes are more directional in nature than those utilized in this study. Hypothesis 1 was not rejected as significant differences occurred in only one of the four comparisons ( $p < .05$ ).

TABLE 3

## SIGNIFICANCE LEVELS OF INDIVIDUAL CORRELATION DIFFERENCES

Time Period	Model IA	Model IM	Model IS
Model 2	N.S. <sup>c</sup>		
IM 3	(IM) <sup>a</sup> .001 <sup>b</sup>		
IM 4	N.S.		
IM 5	(IM) .08		
Model 2	(IS) .001	(IS) .01	
IS 3	(IS) .001	N.S.	
IS 4	(IS) .001	(IS) .01	
IS 5	(IS) .001	(IS) .001	
Model 2	(IA) .001	(IM) .001	(IS) .001
CONJ 3	(IA) .001	(IM) .001	(IS) .001
CONJ 4	(IA) .001	(IM) .001	(IS) .001
CONJ 5	(IA) .001	(IM) .001	(IS) .001

<sup>a</sup>Symbol in parenthesis indicates model which performed significantly better.

<sup>b</sup>Level of significance.

<sup>c</sup>Not significant at the .10 level.

Model IS outperformed model IA ( $p < .001$ ) and model IM ( $p < .01$  and  $p < .001$ ). Hypothesis 2 was rejected.

The performance of model CONJ was quite poor. In fact, all other models tested outperformed model CONJ ( $p < .001$ ). Hypothesis 3 was not rejected. Table 2 shows that the mean attitude-preference rank order correlation using this model varied from .588 to .619, whereas the selected ideal point model correlations varied from .745 to .775. Wright found that respondents viewed this model as if it was an optimization model. The conjunctive model used here performed poorly when used as if it was an optimization model.

## Conclusion

The selected ideal point model outperformed all models evaluated. These were a model excluding the measured ideal point (i.e., the assumed ideal point model found superior in previous research) and the measured ideal point model. For this product category, fast food restaurants, the predictive ability of the multi-attribute attitude model was significantly improved when measured ideal point ratings were selectively included in the model, based on whether the attributes were directional in nature. Ideal point ratings were measured with belief ratings along attributes of the eight brands. The conjunctive model also performed poorly in an optimizing role when used to yield a rank order of brands which was correlated with rank order preference. There is a need to test these models with the criterion measure of probability of choice as use of the conjunctive model leads to the selection of perhaps several acceptable brands rather than a rank ordering of preferred brands.

The aggregate approach to the evaluation of consumer judgment strategies used in this research did not give insight into the distinct and important groups of "minority" consumers who exhibited use of alternative attitude models. For one group of consumers (20 percent of the sample), the conjunctive model yielded the highest correlation with rank order preference. Further research should be conducted using evaluation methods which isolate groups of consumers who use different judgment strategies.

The concept of the ideal point has proved quite useful in the interpretation of attitude structure models: e.g., multidimensional scaling. The concept of the ideal point has major relevance to the marketing manager, and this study has provided a circumstance and a model which used the explicit ideal point measure to increase the predictive performance of the multi-attribute attitude model.

## References

- James R. Bettman, "To Add Importance or Not to Add Importance: That is the Question," Working Paper No. 5, Centre for Marketing Studies, University of California, Los Angeles, October, 1973.
- James L. Ginter, "Attitude Change and Choice Behavior in New Product Introduction," Unpublished Doctoral Dissertation, Purdue University, 1972.
- James L. Ginter, "An Experimental Investigation of Attitude Change and Choice of a New Brand," *Journal of Marketing Research*, 11(February, 1974),30-40.
- Donald R. Lehmann, "Television Show Preference: Application of a Choice Model," *Journal of Marketing Research*, 8(February, 1971),47-55.
- Bernard Pras, "Predictive Qualities of Linear and Non-linear Evaluation Process Models," Unpublished Doctoral Dissertation, Indiana University, 1973.
- Jagdish N. Sheth and W. Wayne Talarzyk, "Perceived Instrumentality and Value Importance as Determinants of Attitudes," *Journal of Marketing Research*, 9(February, 1972),6-9.
- William L. Wilkie and Edgar A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," *Journal of Marketing Research*, 10(November, 1973),428-41.

William L. Wilkie and Rolf P. Weinreich, "Effects of the Number and Type of Attributes Included in an Attitude Model: More is Not Better," Proceedings, Third Annual Conference, Association for Consumer Research, 1972, 325-40.

Frederick W. Winter, "A Laboratory Experimental Study of the Dynamics of Attitude and Choice Behavior," Unpublished Doctoral Dissertation, Purdue University, 1972.

Peter Wright, "Consumer Choice Strategies: Simplifying vs. Optimizing," Journal of Marketing Research, 12 (February, 1975), 60-7.



AN APPLICATION OF MULTIDIMENSIONAL SCALING AND RELATED  
TECHNIQUES TO THE EVALUATION OF A NEW PRODUCT CONCEPT

Larry Percy, Ketchum, MacLeod & Grove

Abstract

A new product concept is studied to determine how it will be received in relation to existing alternatives and whom consumers perceive the likeliest user. Various multidimensional scaling techniques are utilized to analyze the new product concept and seven alternatives as well as five homemaker characterizations and five life-cycle stages.

Introduction

One of the realities of marketing is that more likely than not, consumer researchers are asked to determine the viability of some new product idea already developed by a company's research and development section, rather than asked to help determine what will be a viable new product idea. The need for the study reported in this paper arose from the former.

A new concept for a convenience dinner had been originated, then formulated, by a major food manufacturer. As management assessed the potential of the new idea, they were worried by two vexing questions: 1) How will this product be perceived in relation to existing alternatives; and 2) Who is perceived by consumers as the likely user. The first of these questions was doubly confounding, because if the product was perceived to be similar to existing TV-type dinners, the market would be tough to crack; yet there was much in the new product concept that would suggest it was little more than another TV-type dinner.

A study was designed in which one could evaluate the new product concept within the consumer's product-market cognitive domain, and to examine how the new product concept would be evaluated within various usage scenarios. A multidimensional scaling approach is used in order to fully understand the many interrelationships active in consumer positioning of this new product concept. This multidimensional scaling approach is selected over more conventional concept testing procedures (e.g. a straight "concept-to-use" test) because it is felt what is most important is how the "idea" of this new product will be encoded by consumers for evaluation against alternatives within the product market.

Initial exposure to the new product, assuming it is marketed, is in large part due to advertising and promotion. The image conveyed through these communications variables will be compared and evaluated against competing cognitive stimuli, with evaluation of the new product proceeding from this reference. Establishing this point of reference is thus more critical as a first step toward a decision to market the product: not obviating the need for assessing actual product performance through usage testing, but preceding it.

Study Design

Answers to the two questions asked by management are considered the goals of this investigation. A number of dinner alternatives are considered, along with

several different perceived usage variables. Personal interviews among thirty female heads-of-household in three geographically diverse cities (Boston, Tampa, Omaha) were conducted: 90 completed questionnaires were used in the analysis. Subjects were selected on an area-wide cluster basis, stratified to ensure representation of a broad demographic base.

Method

Two basic exercises were conducted: 1) a gathering of pairwise proximities data among selected dinner stimuli; and 2) a rank ordering of perceived likelihood of serving the selected dinner stimuli under various usage scenarios. In the first exercise, seven alternative dinner options available from the product-market were selected. These are listed below in Table 1. Each of these alternatives was then reduced to a concept statement: e.g. TV Dinner Plus became "frozen dinner featuring family favorites in a single tray with larger portions of the main dish." Along with the new product concept, 28 cards were prepared, each card containing one of the  $n(n-1)/2$  pairs of dinner concepts.

TABLE 1

Selected Alternative Dinner Options

---

---

Box Dinner
TV Dinner
TV Dinner Plus
Main Dish/Starch Combination
Easy to Prepare Homemade Dinner
Quick Homemade Dinner
Big Homemade Dinner

---

---

Subjects were asked to look through all of the cards (which they were told contained the description of two different kinds of dinners), placing all the cards they felt contained descriptions of similar dinners into one pile and all those cards they felt contained descriptions of dinners which were not similar into a second pile. They were then asked to rank the first pile from the card containing the description of the two most similar dinners, and to rank those cards in the second pile from those containing the most dissimilar kinds of dinners. The order of the second pile was then inverted and added to the first, providing a ranking of all 28 pairs from the most to least alike.

In the second exercise, the subjects were presented with a card listing each of the eight concepts used in the first exercise. She was then asked: suppose you were told that a homemaker was described as "modern" i.e. she has a generally contemporary life style without being "hip" or fadish. If this were all you knew about this person, which of these different kinds of dinners do you think she would serve most often? All dinners were similarly ranked, the last being the one the subject perceived this scenario homemaker least often serving. All eight dinner concepts were ranked in a similar manner for the four other homemaker

characterizations listed in Table 2, as well as for the five life-cycle stages shown. These data then provided rank orderings of the perceived likelihood of each dinner being served often by the person described in each scenario.

TABLE 2  
Usage Scenarios

Homemaker Characterizations

1. "modern" women: a generally contemporary life-style without being "hip" or fadish
2. constantly busy around the house
3. unwilling to work or take on outside responsibilities which take her away from her home and children
4. "working" women: a truly active woman who would rather be working than at home all day
5. homemaking/cooking part of wifely role: sees these activities as part of her feminine, wifely role and duty

Life-cycle Stages

1. Young single student
2. Young single job holder
3. Early marriage (no children)
4. Middle marriage (children)
5. Older adult (no children)

Analytic Design

Data collected in the first exercise are considered proximities measurements, and are submitted to Kruskal's (1969) M-D-SCAL 5-M program for non-metric multidimensional scaling. Solutions are sought in three, two, and one dimensions. Data collected in the second phase are considered as two separate data sets for non-metric multidimensional unfolding: the five homemaker characterizations are treated as one set, the five life-cycle stages a second. Spearman rank order correlations are computed for each set in order to determine the "uniqueness" of the rank order profiles. The rank order profiles in each set are then treated as similarity estimates between the eight dinner stimuli and the scenario (cf Percy, 1975). Utilizing again Kruskal's M-D-SCAL 5-M program, this time exercising the lower corner matrix and split-by-rows options, a non-metric unfolding solution is effected for each set in three, two, and one dimension.

Finally, utilizing the appropriate solution configuration from the multidimensional scaling of the eight dinner stimuli, and the rank order profiles collected for each usage scenario, the perceived usage is integrated with the cognitive evaluation via Chang and Carroll's (1969) PROFIT.

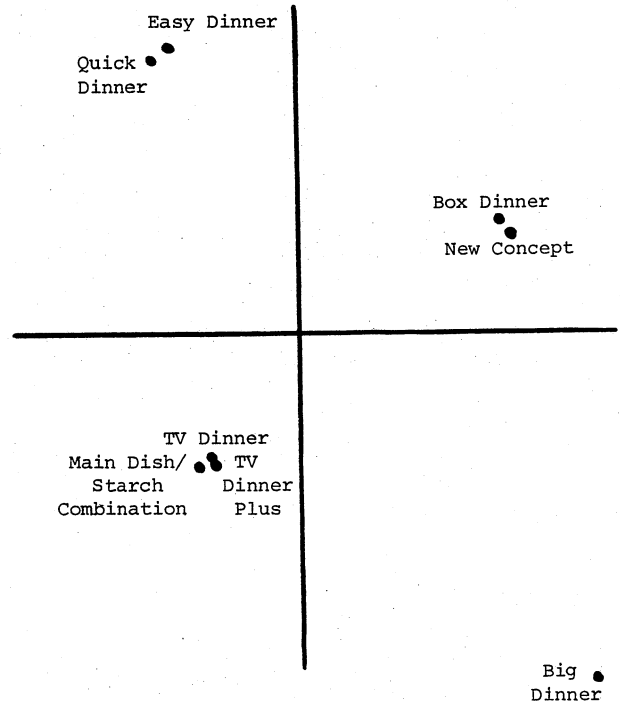
Study Results

The first step in the analysis is to develop an idea of where the new product concept lies in relation to existing alternatives within consumer cognition. Results of the non-metric multidimensional scaling analysis of the  $n(n-1)/2$  pair-wise similarities rankings provide a highly acceptable stress in two dimensions of 0.0097. In fact, the remarkable stress value suggests the possibility of a degenerate or quasidegeneracy of the solution.

Looking closely at the solution configuration shown in Figure 1, one does notice the tendency of the various concepts to collapse into vertices of a parallelogram, one of Shepard's (1974) nine forms of two-dimensional degeneracy. Still, the configuration need not be considered too degenerate for our purposes.

FIGURE 1  
Two Dimensional Multidimensional Scaling  
Configuration of Eight Dinner Concepts

STRESS = 0.0097



The analysis has provided us with a representation of the stimuli in the consumer's mind. What it has perhaps failed to provide is a portion of the actual structural information conveyed by a more non-degenerate solution. It should be noted however that this was an accepted risk in the analytic design, utilizing as it did fewer than ten stimuli for an anticipated two-dimensional solution as well as the inclusion of obviously clusterable stimuli. In fact,

one of the principle goals of the analysis was to determine if indeed the new concept conceptually clustered with the more TV-type dinner stimuli. We have learned that the stimuli do cluster together as anticipated, even though we learned nothing about the relationships among the stimuli within the clusters. But this wasn't important to our purpose: what was important is that we discovered the new concept tended to cluster with box dinners, not TV-type dinners.

Reviewing the resulting multidimensional scaling configuration suggests that four distinct clusters of these stimuli occur within the general consumer cognition, and that these four clusters are considered unlike each other. Looking at the composition of the clusters confirms the logic of the groupings: low vs. high involvement "from scratch" dinners and TV-type vs. other prepared-type dinners. So, while from a statistical standpoint the degeneracy apparent in the solution may be unappealing, from a practical standpoint the solution positions the new concept within existing perception, responding to the stated goal of the analytic design.

The second step in the analysis is to relate the various usage scenarios to perceived appropriateness of the stimuli. The homemaker characterization and life-cycle stage scenario sets were each submitted to a non-metric multidimensional unfolding analysis. The Spearman rank order correlations shown below in Table 3 for the five homemaker characterizations reveal some strong similarities among the profiles, suggesting that a multidimensional unfolding of these data might be difficult.<sup>1</sup>

TABLE 3

Spearman Rank Order Correlations of Homemaker Characterization Perceived Likelihood of Serving Eight Dinner Concepts

Homemaker Characterization	1.	2.	3.	4.	5.
1. "Modern" women	-	.81	.45	.90	.60
2. Constantly busy around house		-	.12	.83	.26
3. Unwilling to work			-	.24	.90
4. "Working" women				-	.29
5. Homemaking/cooking part of wifely role					-

And in fact, the two-space solution configuration revealed in Figure 2 does exhibit a rather high stress value of 0.3536. As a result, one must be careful in

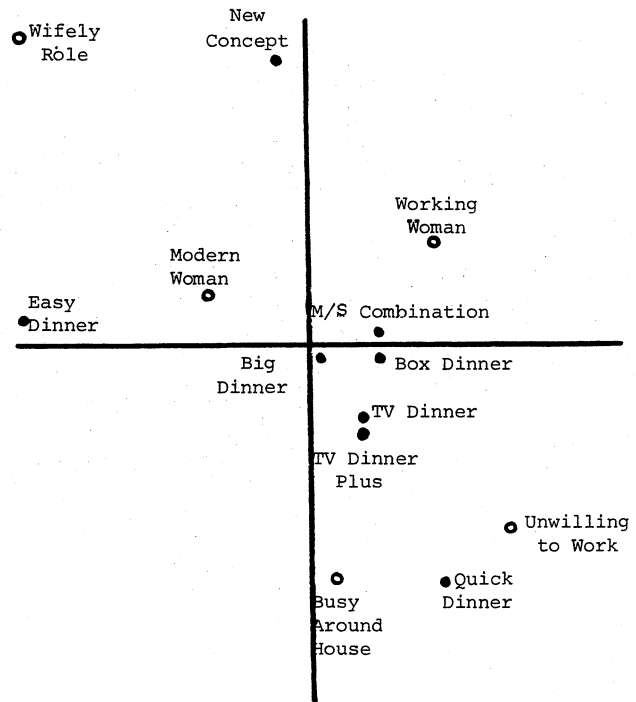
<sup>1</sup>For a more detailed discussion of what to look for in your data to increase the likelihood of a successful multidimensional unfolding, see Moskowitz (1973) and Percy (1975).

drawing any conclusions from the configuration. The fact that five of the stimuli cluster together toward the center of the space suggests that substantially more variation probably occurred between those stimuli. Otherwise, one should expect these five stimuli (the Big Dinner, Box Dinner, M/S Combination, TV Dinner, and TV Dinner Plus) to distribute throughout the space just as the Homemaker Characterization scenarios did.

FIGURE 2

Unfolded Homemaker Characterization Profiles: Two-Space Configuration

STRESS = 0.3536



Looking next at the Life-Cycle Stage Scenarios, the Spearman rank order correlation matrix shown in Table 4 indicates very little similarity among the profiles, unfolding of these data should result in a meaningful solution.

Once again, the early look at the profile data predicted the strength of the solution. Looking at Figure 3, one finds a two-space configuration developed with a stress value of 0.0097. One may certainly consider the inter- and intra-relationships among and between the stimuli and scenarios as meaningful. It is interesting to reflect, for example, on the proximity of the new concept to the Easy Dinner and not the Box Dinner (recall that in the multidimensional scaling of pairwise dissimilarities, the new concept and Box Dinner clustered together in solution). What this perhaps reveals is that in terms of perceived usage, the new concept is rated more constantly like the Easy Dinner

than the Box Dinner, even though it is classified in cognition more like the Box Dinner. In other words, different dimensional criteria are involved with each exercise. (This question is addressed in the final step.) Drawing concentric iso-preference curves from each scenario point reflects the perceived likelihood of each dinner stimuli being served. With the exception of Middle Marriage, the new concept enjoys a rather strong likelihood of being served, especially in Early Marriage and by Older Adults -- and one notes here a constant of no children.

TABLE 4

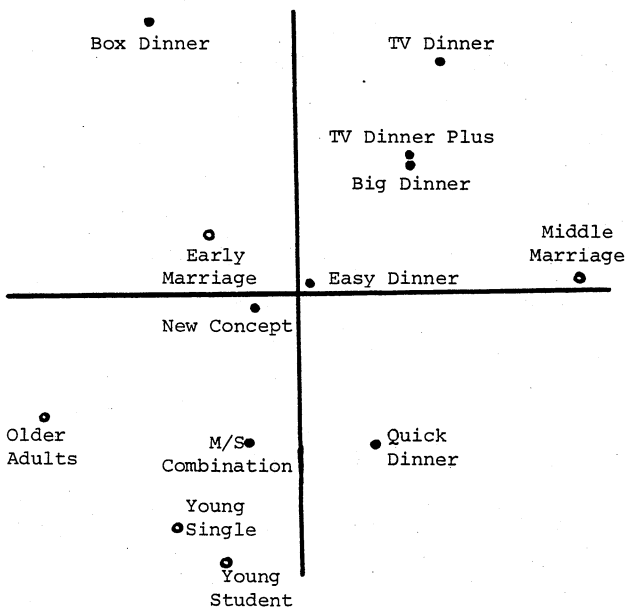
Spearman Rank Order Correlations of Life-Cycle Stage Perceived Likelihood of Serving Eight Dinner Concepts

Life-Cycle Stage	1.	2.	3.	4.	5.
1. Young single student	-	.95	-.57	-.29	-.93
2. Young single job holder		-	-.45	-.48	-.80
3. Early marriage (no children)			-	-.45	.64
4. Middle marriage (children)				-	.17
5. Older Adult (no children)					-

FIGURE 3

Unfolded Life-Cycle Stage Profiles: Two-Space Configuration

STRESS = 0.0097



While the unfolding of the homemaker characterization scenarios did not prove particularly helpful, the unfolding of the life-cycle stage scenarios was revealing of several interesting relationships bearing on a positional evaluation of the new product idea.

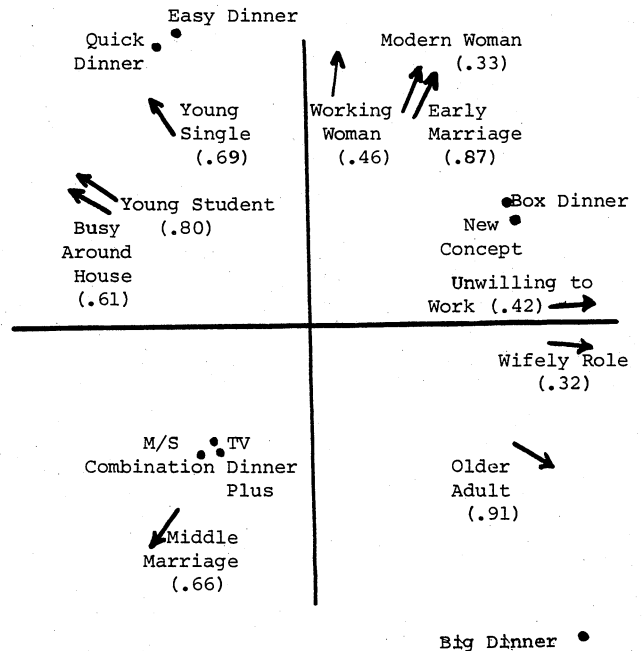
The final step in the analysis is the integrating of the perceived usage information with the general cognitive representation. Each of the rank order scenarios were introduced as properties, along with the coordinate values of the eight dinner stimuli as determined by the M-D-SCAL solution in two dimensions, into Chang and Carroll's PROFIT algorithm. The results provide a vector determination, or direction, for each property (or scenario in this case) in the original two dimensional space such that the projections of the eight dinner concepts on that vector correspond optimally to the given rank order of the perceived likelihood of it being served under that scenario.

Looking now at Figure 4, one finds that life-cycle stage scenarios enjoy good fits to the cognitive representation, while the homemaker characterizations (with the exception of Busy Around House) offer poor fits. This conclusion is drawn by examining the Rho values for each property vector enclosed in parentheses, representing the maximum correlation between each given property and its fitted vector.

If one ignores the homemaker characteristic scenarios, owing to their poor fits, the new concept finds itself associated quite highly with the Early Marriage and Older Adult life-cycle stage scenarios. It would seem that our new product idea tends not to be associated in consumers' minds with the TV-type dinners, which are more oriented in perceived usage to Middle Marriage; rather it is perceived to have a higher likelihood of service among married homemakers with no children, and to cluster with Box Dinners.

FIGURE 4

10 Property Vectors Fitted in the Average-Subject Stimulus Space of 8 Dinner Concepts Using Profit



## Discussion and Conclusion

In this study, attention has been drawn to the application of multidimensional scaling and related techniques in the evaluation of a new product concept. The focus is purposefully narrow, owing to an overriding concern on the part of management that the new product concept may be perceived by consumers as similar to TV-type dinners. The application of multidimensional scaling procedures rather than more conventional concept testing methods was suggested by this need to determine where the new product concept would be situated within consumer cognition. They are proven particularly useful and sensitive in uncovering certain aspects of cognition, and provide a graphic representation.

Building on past research in the area of food and the meal, a number of made-from-scratch and packaged-convenience dinner alternatives were reduced to simple concept form for comparison with the new product concept. The logic here holds that the introduction of the new product will rely heavily on the consumer's understanding of the "concept" once aware of it, and subsequent trial and usage will be critically mediated by comparisons of her perception of the new product versus competing alternatives. As a further aid in gaining a complete understanding of her cognitive associations, specific homemaker characteristic and life-cycle stage scenarios were considered for all competing dinner alternatives (including the new product concept).

The multidimensional scaling of the pair-wise concept similarities, despite a possible mathematically degenerate configuration, clearly indicated that consumers would not encode the new product concept together with TV-type dinner alternatives. This was, of course, welcome news from a marketing standpoint.

Multidimensional unfolding of the concepts and scenarios indicated that women apparently find it difficult to form any consensus regarding possible dinner alternatives and particular homemaker characterizations. The very high resulting stress measure (presaged by high profile correlations) reflected a great deal of similarity in perceived usage attribution. While women may in fact expect little meal serving difference between the homemakers characterized by the scenarios presented, it may also be true that they found the characterizations difficult to relate to this type of behavior.

On the other hand, differences were easily related to life-cycle stages. The multidimensional unfolding in this case revealed meaningful associations between specific life-cycle stages and the likelihood of often serving particular dinner alternatives. The new product concept was related strongly to the early married and older adult stage (both characterized by no children).

A final exercise combining the cognitive representation with the scenario attribution via PROFIT reinforced the results of the unfoldings. Once again very poor fits generally were found for the homemaker characteristic scenarios, but strong fits for the life-cycle stage scenarios. This embedding of the scenarios in the multidimensional scaling solution represented an external analysis verification of the internal analysis evaluations presented by the multidimensional unfoldings. The consistency of results offers a certain implied level of confidence in the conclusions to be drawn.

We have seen how the utilization of multidimensional scaling and related techniques may be productively applied to new product concept evaluation, particularly when an initial assessment of generalized cognition is important. Having satisfied one's self of the market positioning potential of the general concept, it is possible to proceed comfortably with more detailed analysis (e.g. conjoint measurement of the concept components and possible positionings) evaluating the strength of the concept.

## References

- J. J. Chang and J. D. Carroll, "How to Use PROFIT, A Computer Program for Property Fitting by Optimizing Nonlinear or Linear Correlation," Bell Laboratories, unpublished manuscript, 1969.
- J. B. Kruskal and F. J. Carmone, "How to Use M-D-SCAL, A Program to Do Multidimensional Scaling and Multidimensional Unfolding," (Versions 5M of M-D-SCAL, all in Fortran IV), Murray Hill, N.J.: Mimeo, Bell Telephone Laboratories, 1969.
- H. R. Moskowitz, "Profile Attributes as Similarities," paper presented at Association for Consumer Research Fall Conference, Boston, 1973.
- L. H. Percy, "Multidimensional Unfolding of Profile Data: A Discussion and Illustration with Attention to Badness-of-Fit," *Journal of Marketing Research*, 12 (February, 1975), 93-9.
- R. N. Shepard, "Representation of Structure in Similarity Data: Problems and Prospects," *Psychometrika*, 39 (December, 1974).

LINK IN PREDICTING CONSUMER BEHAVIOR

Larry Percy, Ketchum, MacLeod & Grove

Abstract

The lack of reliable empirical support for the intuitive connection between personality and behavior is discussed. Two explanations are presented to account for this low accountability: a) the use of personality traits rather than profiles; and b) a lack of cognizance of the personality-attitude-behavior link and the actions of intervening variables. Study results are discussed that address these points and indicate strong predictive ability.

Introduction

Despite an apparent lack of success in empirical explanation, personality as a predictor or mediator in communications and buyer behavior theory persists. The nexus is drawn over and over in the literature. Howard and Sheth (1969) describe personality traits as those characteristics that account for differences among people and that are predictive of their behavior (emphasis added). They go on to discuss a number of "significant" ways in which certain personality traits are important in their theory of buyer behavior, as an element of the buyer's frame of reference. Engel, Kollat and Blackwell (1973) talk about personality as the sum total of factors involved in each individual's way of thinking, behaving and responding that make their image. They proceed to devote an entire chapter to the relationship between personality and consumer behavior. Parr (1972) suggests that, other things permitting, the preferences and demands of an individual's personality will affect his choice of surroundings. From this one might easily infer that a person's purchase behavior and media mode selection are components of his surrounding, and thus a function of personality.

The importance of understanding personality from a communications standpoint is mentioned by Howard and Sheth (1969), and discussed extensively by McGuire (1968, 1972) among others. McGuire (1968) spends a great deal of time detailing the effects of specific personality traits (e.g. self-esteem) on the various stages of his informational processing paradigm and notions of attitude change. Particular attention is drawn to the mediational aspects of many variables in understanding the effect of personality on a behavioral or attitudinal change.

Perhaps this is a portent to why it has been so difficult to demonstrate in a straight forward manner a strong relationship between personality and behavior. As McGuire points out, this mediational principle in itself does not allow one to make any confident predictions, only to call our attention to the complexity of the undertaking. And complex it has proven. Typically only about five or ten percent of the total variance in buyer behavior has been accounted for by personality construct measures. But as Kassarian (1971) concludes his review of research in personality and consumer behavior, to expect the influence of personality variables to account for a large portion of the variance is most certainly asking too much. He lays a good deal of this blame at the feet of the many other interrelated influences on the consumer decision

process.

Other explanations for this low accountability have been offered. Nakanishi (1972) finds that both Kassarian and Jacoby (1971) attribute this uncertain predictive ability to either the lack of a proper hypothesis to test or sloppy adaptation of an otherwise clinical measurement of personality. Returning to Engel et al., their feeling is that personality characteristics are better thought of as moderator variables, and hence one should expect a better prediction of behavior in one situation than another. Crane (1972) suggests that the personality of most individuals is so multifaceted that they can meet the demands of any situation they are likely to encounter. He therefore feels role tends to be more useful than personality in predicting and influencing human behavior. Yet one of the four roles he finds important to the marketing communicator (family decision making unit, those of husband or wife) Engel et al. suggest is influenced by the personality characteristics of the individuals involved in decision making. They offer the example of wives who have a strong need for love and affection generally having less influence in purchasing decisions. One seems to be traveling on a Moebius Strip.

A more rational explanation is perhaps to be found in Rosenblatt and Miller (1972): they concern themselves with experimental design. While many researchers evaluate response to some social event (e.g. an advertising message) among people who differ in personality, unfortunately personality is almost always correlated with pre-test differences on the communication topic. Personality becomes confounded with the message or source: And there are scale problems (other than those of the personality measurement instrument). As discussed below, in addition to confounding in communications, there may be confounding in attitude and behavior.

Beyond any possible problems in research design, however it is the purpose of this paper to discount many of the explanations offered above, and to suggest that the reasons for the low accountability and conflicting results to date of research in personality and behavior are: a) reliance on personality traits as opposed to personality profiles; and b) a failure to recognize the importance of attitude or disposition as a mediator of behavior.

Background

Personality implies a whole, not a string of independent personality traits. Why should it be surprising that a single personality trait, which may co-vary in a multitude of ways with other traits, fails to contribute on its own much to the prediction of a specific consumer behavior such as a brand or product choice. It makes far better sense to consider a profile of personality, one that reflects the interactions of various personality traits. As McGuire (1972) has pointed out, any chronic personality trait that tends

to make an individual extremely resistant (say) or extremely receptive to a particular influence probably becomes embedded over time in a matrix of other traits that serve as correctives in moving the person towards an intermediate level of behavior.

Sparks and Tucker (1971), after noting essentially weak or spotty relationships between specific personality traits and the use of particular products, went on to demonstrate significant relationships between profiles of these same traits and bundles of various product usage. Building on these results Alpert (1972) studied the relationship between personality structure and product attributes. Although he cited a few limitations, he none-the-less found strong relationships in matching personality profiles of product attributes' determinance. It seems that at least some support is developing for believing that the use of personality profiles, rather than specific personality traits are more helpful indicators of consumer behavior

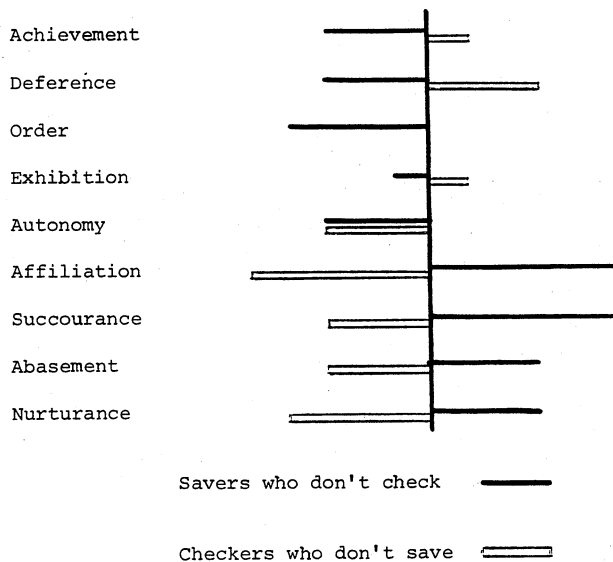
### Utilizing Personality Profiles

In seeking an appropriate instrument for personality measurement, initially an adaption of Edward's EPPS was used. Fifty of the original paired statements were retained after a great deal of testing, providing measures of nine personality traits. At issue was a need to provide an instrument that would not be overly taxing within the context of an otherwise long interview, and one which would enjoy a high level of participation on the part of subjects.

An early application in a study of banking behavior provided a very interesting set of profiles contrasting those who have savings accounts but no checking account vs. those who have checking accounts but no savings account. As detailed in Figure 1 below, one

FIGURE 1

Personality Profiles of Savers Who Don't Check Vs. Checkers Who Don't Save (-) (+)



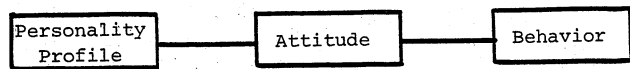
sees that the two groups exhibit almost completely opposite personality profiles. The importance of this finding was critical to the development of advertising copy: in talking to the saver who doesn't check, the creative tone should be soft and the message coddled in a sort-of "we know what's best" manner. Quite the contrary of course would apply to the checker who doesn't save; they would be uninterested in any soft-selling copy. As Engel et al. (1973) suggest, advertising copy can be much more effective if the artists and writers have a rich understanding of the total lifestyle of those to whom they are writing. Here is an example of a strong relationship between personality, expressed as a profile rather than a specific trait, and consumer banking behavior. Understanding this relationship resulted in more effective advertising strategy.

### Relating Personality, Attitude and Consumer Behavior

Attention is now turned to an examination of personality (expressed as a profile) as a common denominator among consumers who tend to hold similar attitudes toward a product-market, and in relating them, through these commonly held attitudes to specific behavior as shown in Figure 2. Personality is considered a determinant of behavior, but only through an underlying attitudinal construct. Personality relates to commonly held attitudes and perceived behavioral orientations, and it is these factors that mediate specific behavior.

FIGURE 2

### The Personality-Attitude-Behavior Link



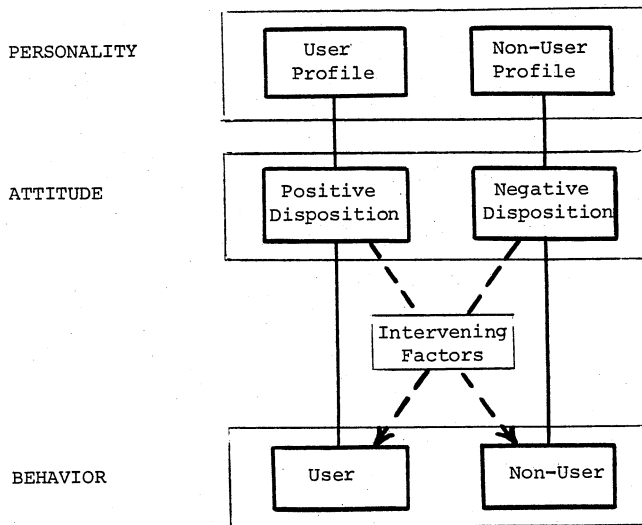
Being able to describe a target market group in terms of their gross personality characteristics, which are in turn known to be good discriminators of a particular attitude or disposition, will provide a much clearer picture of whom one is dealing with for purposes of developing advertising and marketing strategies directed toward them. This connection with attitudinal disposition is important from a strategic viewpoint, for it is attitude that the marketer or advertiser seeks to exploit, change or modify in effecting positive behavior (i.e. usage and trial) toward his product.

However, it must be recognized that even given the above chain, intervening factors may effect behavior contrary to that predicted by a given attitudinal disposition. As an example, imagine the man who hates lawn work, but finds himself with an unruly hedge on his hands. He may reluctantly purchase a hedge trimmer even though his personality profile matches that of a non-owner of lawn and garden tools. This fact of ownership does not make him a good prospect for additional lawn and garden tools. The intervening factors of an overgrown hedge and complaining neighbors effected his behavior. Figure 3 illustrates the possible action of intervening factors.

Two studies are discussed below: the first illustrating the Personality-Attitude-Behavior link; a second the results of intervening factors.

FIGURE 3

A Diagram of the Personality-Attitude-Behavior Link  
And the Action of Intervening Factors



In the first study, a large scale inquiry into the meal among a nationally projectable sample of over 1,200 homemakers, personality measurements via the modified EPPS were gathered in addition to attitudinal and behavioral measures. The attitudinal data, gathered from five-point Likert scale measurements of 20 statements (which were developed from earlier research on a much larger list), provided the basis of an attitudinal segmentation of the market. Following Cattell (1966a, b) a Q-mode factor analysis revealed six groups of homemakers with different attitudes toward the meal. Personality profiles over nine traits were developed for each group.

These profiles were intercorrelated for each of the attitude groups, with the results reported in Table 1.

TABLE 1

Intercorrelations of Personality Profiles  
For Six Meal Attitude Groups

Attitude Group	Q <sub>1</sub>	Q <sub>2</sub>	Q <sub>3</sub>	Q <sub>4</sub>	Q <sub>5</sub>	Q <sub>6</sub>
Q <sub>1</sub>	-					
Q <sub>2</sub>	-.536	-				
Q <sub>3</sub>	.769	-.623	-			
Q <sub>4</sub>	-.382	.713	-.388	-		
Q <sub>5</sub>	-.050	-.692	-.059	-.529	-	
Q <sub>6</sub>	-.671	.546	-.545	.561	-.049	-

One notices a good correlation between attitude groups Q<sub>1</sub> and Q<sub>3</sub> of .769 and between attitude groups Q<sub>2</sub> and Q<sub>4</sub> with .713 (and possibly Q<sub>6</sub>). The thesis of this paper would suggest that because the personality profiles of these groups are similar (i.e. correlate well), this could be predictive of a common underlying attitude. In fact, this is precisely what one finds. The first "cluster" of Q<sub>1</sub> and Q<sub>3</sub> reflect two groups of homemakers whose attitudes might be described as traditional, centering on a consumer attitude of heavy personal involvement with the preparation and serving of the meal. They enjoy spending time in the kitchen. The second "cluster" of Q<sub>2</sub> and Q<sub>4</sub> express a strong interest in convenience foods, and little desire to be heavily involved in the meal. The positive correlations with this "cluster" by Q<sub>6</sub> may be explained by their lack of interest in meal involvement; however, it is not as strongly correlated for they do not share the same attitude toward convenience foods.

There are clearly strong personality-attitude relationships evident in these data to support the first premiss, but what of the next link, to specific behavior. While the attitudes measured are perhaps a little too general to expect brand usage associations, Table 2 none-the-less illustrates a good relationship between the "traditional" groups (Q<sub>1</sub> and Q<sub>3</sub>) and non-users of the brand studied. This makes sense when it is understood that the brand in question is identified with main meal entree convenience foods, something a real "traditional" homemaker would be wont to use. While some associations might have been expected among the other attitude groups and brand users, it should be pointed out that only about five percent of all homemakers are users of this particular brand. The disposition to purchase may be positive, but marketing realities have thus far precluded it. It would appear that the personality-to-attitude-to-behavior chain enjoys a certain level of credibility.

TABLE 2

Personality Profile Correlation Matrix of Attitude  
Groups and Brand Users/Non-Users

Brand Usage	Q <sub>1</sub>	Q <sub>2</sub>	Q <sub>3</sub>	Q <sub>4</sub>	Q <sub>5</sub>	Q <sub>6</sub>
Use	-.753	-.003	-.583	-.192	.432	.181
Never Use	.812	-.125	.664	.131	-.393	-.264

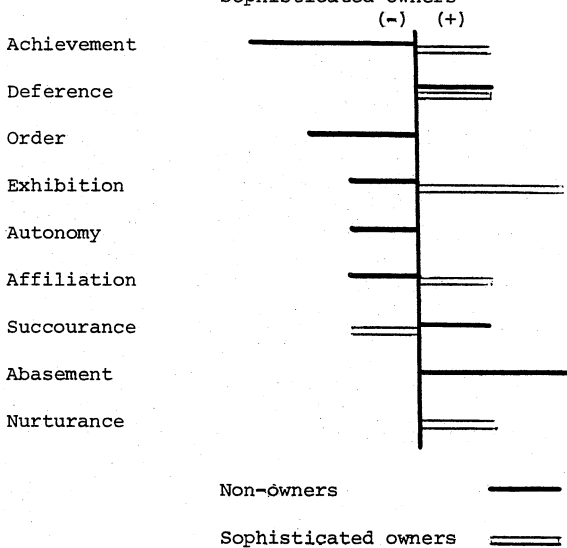
In a second study both attitudinal groups and groups clustered on the basis of perceived behavior were profiled and compared with specific product usage groups. Unlike the above study, these data were collected from a national probability sample of men, and a modified version of the Gough and Heilbrun Adjective Check List (1965) replaced the modified EPPS. The nine scales utilized in the EPPS earlier are retained, along with the addition of six scales. Reasons for the instrument change centered largely on the reduction of administration time from about twenty minutes to five with the ACL with no apparent loss in discrimination.



The product category studied was light durables of male interest where ownership is known to be widely distributed over a number of different products, with individual subjects likely to own none, one, a few, or many items, each item available in a broad range of price, quality and sophistication. Examining the personality profiles between non-owners and sophisticated owners (so-called because of multiple product ownership coupled with ownership of more sophisticated items) revealed opposite profiles (correlating  $-.595$ ). These profiles are displayed below in Figure 4 for the nine personality traits comparable with the EPPS profiles shown earlier.

FIGURE 4

Personality Profiles of Non-Owners Vs. Sophisticated Owners



Interpreting these profiles, one could say that those who own no items in this category may be described as unsure of themselves and their abilities, seeking stability and continuity in their environment, and apprehensive of ill-defined or risk-involving situations. They try to avoid situations calling for choice and decision-making, and are dubious about the results of expending effort or becoming involved with their labor. On the other hand, the sophisticated owners are determined to do well, and are usually successful. They have a quiet confidence in their own abilities and worth, welcoming the challenges to be found in disorder and complexity. Sincere and dependable, they are adaptable, resourceful, and comprehend problems and situations rapidly and incisively.

There is clearly a relationship, and a significant one, between the two extremes of category behavior and personality profiles. The study goes on to examine the predictive ability of these profiles in correctly classifying ownership with and without a mediator.

Personality profiles were developed for attitudinal groups generated by a Q-mode factor analysis of 36 attitude statements and the clusters generalized by the Singleton-Kautz (1969) minimized within cluster variance algorithm on 26 perceived behavior statements, and the results correlated with the above profiles.

These results along with the incidence of non-owners and sophisticated owners in each cluster and Q-group are shown in Table 3. While the personality profile correlation between non-owners and Cluster 1 is strong (.933), predicting the heavy incidence of non-owners in that group (81%), equally strong correlations of non-owners with Cluster 4 and Q-groups 3 and 4 (.899, .656, and .917) would also suggest heavy incidences of non-owners in these groups, when in fact non-ownership is quite low (9%, 12%, and 33%).

TABLE 3

Ownership Penetration and Personality Profile Correlations Among Behavioral Clusters And Attitudinal Groups

	Non-Owner	
	Percent of Group	Personality Profile Correlation
Cluster 1	81%	.933
2	7	-.874
3	14	-.721
4	9	.899
5	4	.114
Q-Group 1	23%	-.180
2	7	-.936
3	12	.656
4	33	.917
5	15	.279
	Sophisticated Owner	
	Percent of Group	Personality Profile Correlation
Cluster 1	2%	-.645
2	58	.841
3	26	.215
4	24	-.621
5	18	-.514
Q-Group 1	22%	-.510
2	45	.783
3	20	-.029
4	18	.469
5	50	-.369

This seems to debunk the notion of the predictive ability of personality profiles that has underlain a great deal of this paper. However, this is only the case if one fails to address the question of intervening factors as a mediator in the predictive process. In the detailed findings of this study reported elsewhere (Percy, 1975), the cluster and Q-groups were themselves clustered together according to similarity of personality profile and considered for compatibility of meaning. The results provide a three step construction representing negative behavior and attitude within the product market (P<sub>1</sub>), positive behavior and attitude (P<sub>2</sub>), and a benign group (P<sub>3</sub>). This accounted for 100%

of the behavior component and 90% of the attitude component under study.

Table 4 details the personality profile correlations of these three groupings with non-owners and sophisticated owners. It is clear that the personality profile of non-owners is strongly related to that of P<sub>1</sub>, the negative behavior and attitude grouping. Each of the four components of this grouping reflected negative attitude toward usage, even though for other reasons certain of these subjects none-the-less own some items from the category. This is unimportant, for this usage cannot be construed as positively motivated -- their attitude and motivation is against usage, and it is this disposition which the personality profiles have successfully predicted. The same follows for the relationship between the sophisticated owners and P<sub>2</sub>. It would appear that personality, although not a unique predictor of ownership, still discriminates well the underlying behavior and attitude factors that originate ownership.

TABLE 4  
Average Correlations of Personality Profiles

Personality Profile Clusters	Sophisticated	
	Non-Owners	Owners
P <sub>1</sub> = (C <sub>1</sub> , C <sub>4</sub> , Q <sub>3</sub> , Q <sub>4</sub> )	.851	-.441
Cluster 1	.933	-.645
Cluster 4	.899	-.721
Q-Group 3	.656	-.029
Q-Group 4	.917	-.469
P <sub>2</sub> = (C <sub>2</sub> , C <sub>3</sub> , Q <sub>2</sub> )	-.844	.613
Cluster 3	-.874	.841
Cluster 3	-.721	.215
Q-Group 2	-.936	.783
P <sub>3</sub> = (C <sub>5</sub> , Q <sub>1</sub> )	-.033	-.512
Cluster 5	.114	-.514
Q-Group 1	-.100	-.510

#### Discussion

The role of personality in the study of consumer behavior is frequently discussed. As we have seen, it often finds its way into the literature, but rarely with any empirical certainty as to its actual association with specific behavior. Numerous explanations have been ventured as to why this low accountability is experienced, yet none seem satisfying. This paper advances a two-fold explanation, illustrated and supported by specific examples.

First, personality traits are indicated to be less effective than personality profiles in predicting specific consumer behavior. Evidence is cited from Sparks and Tucker (1971) and Alpert (1972) showing the increased predictive power of the personality profile. An example of strong discrimination in banking behavior is presented, illustrating where personality

profiles offered critical guidance in the development of targeted advertising strategy.

Second, and more importantly, it is suggested that a personality-attitude-behavior link exists, and that while personality (expressed as a profile) is a consistent predictor of attitudinal disposition, intervening factors often effect actual behavior in a manner contrary to that suggested by an individual's personality and attitude. Two studies confirm the relationship between commonly held attitudes toward a product-market and personality; one continuing the chain and predicting brand behavior, the other demonstrating the effect of intervening factors on prediction.

From an advertising or marketing strategic viewpoint, it should be pointed out, the prediction of a specific behavior (e.g. Brand choice) is not as important as the prediction of a target group's attitudinal disposition. What one is generally interested in is potential future behavior, usually as a consequence of some advertising or marketing effort. To the extent that one may use personality to better understand a target group holding a particular set of attitudes (selected because their attitude presages a disposition toward positive behavior), one should be in a better position to more effectively communicate to that group. One is maximizing potential through the exploitation, modification or changing of specific attitudes in one's advertising and marketing efforts. This personality-attitude-behavior link plays an important part in understanding the best way to go about it.

In the past, personality has been less than successful in predicting specific consumer behavior because: a) it has used traits rather than profiles; and b) it did not take cognizance of the personality-attitude-behavior link and the actions of intervening factors. This analysis does not obviate the intuitive supposition that behavior should be predictable knowing an individual's personality; rather it offers an explanation of the connection.

#### References

- M. I. Alpert, "Personality and the Determinants of Product Choice," Journal of Marketing Research, Vol. IX (February, 1972) 89-92.
- R. B. Cattell, "The Data Box," in R. B. Cattell, ed. Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally Company, 1966a.
- \_\_\_\_\_, "The Meaning and Strategic Use of Factor Analysis," in R. B. Cattell, ed. Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally Company, 1966b.
- E. Crane, Marketing Communication 2nd ed. New York: John Wiley & Son, 1972.
- J. F. Engel, D. T. Kollat and R. D. Blackwell, Consumer Behavior 2nd ed. New York: Holt, Rinehart & Wilson, 1973.
- H. G. Gough, A. B. Heilbrun, Jr., The Adjective Checklist Manual. Palo Alto, California: Consulting Psychologists Press, 1965.
- J. A. Howard, and J. N. Sheth, The Theory of Buyer Behavior. New York: John Wiley & Son, 1969.

J. Jacoby, "Personality and Innovation Proneness," Journal of Marketing Research, Vol. VIII (May, 1971) 244-47.

H. H. Kassarjian, "Personality and Consumer Behavior: A Review." Journal of Marketing Research, Vol. VIII (November, 1971) 409-18

W. J. McGuire, "Attitude Change: The Information-Processing Paradigm," in C. G. McClintock, ed. Experimental Social Psychology. New York: Holt, Rinehart & Winston, 1972.

\_\_\_\_\_, "The Nature of Attitudes and Attitude Change," in G. Lindzey and E. Aronson, eds. Handbook of Social Psychology Vol. 3. Reading, Mass. 1968.

M. Nakanishi, "Personality and Consumer Behavior: Extensions," Proceedings. Third Annual Conference Association for Consumer Research, 1972, 61-65.

A. E. Parr, "In Search of Theory," in H. M. Proshansky, W. H. Iттerson and L. G. Rivlin, eds. Environmental Psychology. New York: Holt, Rinehart & Winston, 1970.

L. Percy, "Relating Personality Profiles and Consumer Behavior via Underlying Attitude and Perceived Behavior" paper presented at the American Psychological Association 83rd Annual Convention, Chicago, 1975.

R. C. Singleton, "A Fortran Program for Minimum Variance Clustering of Multivariate Data," Mathematics and Statistics Division Memorandum, Office of Naval Research, October, 1969.

D. L. Sparks and W. T. Tucker, "A Multivariate Analysis of Personality and Product Use," Journal of Marketing Research, Vol. VIII (February, 1971) 68-70.

AN EXPERIMENTAL INVESTIGATION OF SITUATIONAL  
EFFECTS ON RISK PERCEPTION

Mark Vincent, Oklahoma State University  
William G. Zikmund, Oklahoma State University

Abstract

This exploratory study was designed to determine the effects of different buying situations on several dimensions of perceived risk. Statistical analysis of generated data indicated significant differences in responses to selected risk dimensions due to both situational main effects and situational-interaction effects. These results indicate that buying situations should be considered as partial determinants of perceived risk.

Marketing researchers frequently have suggested that their results have been influenced by "situation-specific" variables. While it has long been held that situational factors would aid substantially in explaining the variability which exists in buyer behavior, only recently have consumer psychologists attempted to systematically investigate the influence situational variables have on brand or product choice.

Most recently, Belk (1974) found that situational variables had a significant influence on consumer preference toward meat and snack product classes. Lutz and Kakkar (1975) applied the Mehrabian-Russell Emotional Mediators Instrument to analyze for psychological differences in the same buying situations which Belk considered. Their results basically confirmed the previous results of situational effects and situational-product interaction effects. In Sandell's (1968) research, situational factors such as "for breakfast," "with lunch," "at a party," explained a larger portion of total variance and beverage preference behavior than did Brand. Miller (1975) found a situational multi-attribute attitude model outperformed the traditional (non-situational) model when used to predict overall choice. Thus, the evidence is relatively clear that situational factors, at least marginally, improve prediction of brand choice for specific product categories.

By introducing situational variables into the research design, consumer psychologists should be able to account for more variance than by ignoring or controlling for them (Belk, 1975). However, Lutz and Kakkar hold the percent explained by situation is likely to be of the same magnitude as personality, i.e., "the situation in and of itself is not a powerful predictor of consumer behavior."

The objective of this research is to isolate which cognitive variables situational influence mediates. More specifically, rather than studying "situation" as an isolated construct to explain consumer behavior, our focus is to examine the relationship between situation factors and perceived risk. Hopefully, establishing experimental verification for the theoretical notion that buying situations may have a systematic and demonstrable effect on dimensions of perceived risk.

The Buying Situation

There is no general agreement concerning the definition of a situation. (See Lutz and Kakkar 1975 for a review). A locus of time and space is one of the simplest definitions. Belk (1975) argues that a rich and meaningful use of the concept "situation" includes both

the antecedent conditions for the momentary individual states which the person brings to a given time and place and the physical features which he finds there. An objective situation may be externally verified without measuring intrapersonal factors.

"... objective descriptions may include the existence of external facts and events which bear upon current behavior even though they are not themselves physically a part of that situation."

"For example, if it can be determined that the stimulus object ... is an item to be served at an upcoming dinner party for the new boss, these objective descriptions can be included in the situational specification directly rather than attempt to measure whether or not the situation is regarded to be threatening by the individual."

(Belk, 1975, p. 429)

However, several researchers tend to agree that there is a need to adopt a subjective view of this situation in order to understand how situational factors influence consumer behavior. Holding that individuals transform situational input into behavioral output, Lutz and Kakkar offer a definition that includes explicit consideration of perceptual processes:

The psychological situation, ... may be defined as an individual's internal response to, or interpretation of, all factors particular to a time and place of observation which are not stable intra-individual characteristics or stable environmental characteristics, and which have a demonstrable and systematic effect on the individual's psychological processes and/or his overt behavior.

This definition complements rather than competes with Belk's viewpoint and is adopted for purposes of this research.

Perceived Risk and Buying Situations

Few addresses to the American Marketing Association have stimulated as much empirical research as Bauer's 1960 paper on perceived risk. Perceived risk, a subjective evaluation process, has been associated with new product adoption, brand and store loyalty, mode of shopping as well as word of mouth communications and opinion leadership. (For an excellent review, see Ross, 1974). Studies by Cox (1964) and Spence, Engle, and Blackwell (1970) indicated that perceived risk is affected by factors other than product. Each of these studies found that the place of shopping, a situational factor, influenced perceived risk. Each, however, was concerned with the physical surroundings of buying situation, "where to buy." Buying situations may also be concerned with dimensions such as the interpersonal surroundings (group opinion), time factors, goal structure (cf. Miller, 1975 Lutz and Kakkar, 1975) and may have a profound effect upon perceived risk associated with the product. It is therefore, the intention of this study to test if

specific buying situations, not related to physical surroundings, are related to the various dimensions of perceived risk.

### Perceived Risk

The initial empirical studies related to perceived risk utilized the uni-dimensional measures of overall risk associated with certain products or brands. (e.g., Arndt, 1967, Cunningham, 1965). More recently consumer behavior theorists have recognized the need to identify risk as a multi-dimensional phenomena. General risk has been variously identified into dimensions such as social, performance, physical, and financial risk (e.g., Roselius, 1971; Perry and Hamm, 1969; Zikmund and Scott, 1973). This research views perceived risk as a multi-dimensional phenomena and utilizes the various types of risk as dependent variables.

### Experimental Methodology

A two-by-two factorial experiment was designed to measure the influence of situational factors and store on perceived risk. The experiment was a factorial design with repeated measures on store factors.

#### Situation \* Store

Figure 1 is a representation of this design. In Situation 1, respondents were asked to consider the purchase of an electric knife for home use. In Situation 2 they were asked to consider the purchase of an electric knife to give as a wedding present to a close friend. Store 1 was identified as Montgomery Ward. Store 2 was identified as Gibson's Discount Store.

FIGURE 1

	Store 1	Store 2
Situation 1		
Situation 2		

#### Factorial Design of Experiment

Each member of the sample taken was randomly assigned to a situational treatment and responded to both store types. While extensive pretesting indicated that differences may exist across these situations, the use of these two stores was based upon findings by Hisrich and others (1972) that differences in perceived risk may exist in store selection.

Stores were selected on the basis of availability to members of the sample rather than on any specific a priori distinction. In other words, two general department stores were selected, rather than identifying the stores as "an unknown store" versus a "well known store" or some other type of manipulation.

The order of appearance of the two types of stores was randomized within the measurement instrument. The home usage situation and the gift situation were intended to elicit differences which may exist between factors considered in buying for one's self and buying for another person i.e., the situation dimensions, goal structure and interpersonal opinion were manipulated (cf. Lutz and Kakkar, 1975 p.3).

This experiment was structured to provide information concerning situational effects, store type effects, and the interaction effects between store type and situation

on perceived risk. It is concerned with a relatively inexpensive consumer durable, a home usage situation, a wedding gift situation, and two types of stores which are both recognized by and available to the intended sample. In this respect, the experiment could be considered to realistically reflect the realm of experience of the intended sample, i.e., it meets the requirement that the situation be encountered frequently (Miller, 1975). Regardless of whether an electric knife had ever been purchased for either of these reasons and/or from either of these store types, it is not unreasonable to assume that such considerations may have been made or could be made.

The selection of situations for both experiments was based upon a pretest of thirty-four respondents, seventeen from Ponca City, Oklahoma, and seventeen from Midwest City, Oklahoma, which indicated that significant situational effects might exist.

### Hypothesis

Just as individual risk dimensions are expected to influence the usage of selected information sources, we expected that certain types of buying situations would be more prone to influence specific dimensions of perceived risk. For example, situations varying in degrees of interpersonal relations would be expected to mediate the amount of social risk perceived by consumers. The general null hypothesis is as follows: There is no relationship between the type of buying situation and the dimensions of perceived risk. Another hypothesis related to stores is as follows: There is no relationship between the type of store which a person perceives and the dimensions of perceived risk.

### Sampling Procedure

A sample of forty-four housewives and wives working outside the home was taken on a voluntary basis through personal contact and door-to-door solicitation. Twenty-two of these subjects were residents of Midwest City, Oklahoma, a suburb of Oklahoma City, and twenty-two of these subjects were residents of Stillwater, Oklahoma. T-tests indicated no significant differences in the overall manner in which the subjects of these two communities responded to the experimental measurement device.

### Data Collection

In order to generate experimental data, each of the forty-four respondents received a self-administered questionnaire which contained one situation with both store types. The assignment of situational treatments, as well as the order of presentation of store type, were randomized so that as many respondents as possible received different orderings and situational treatment pairs.

Each page of the questionnaire contained a brief statement of the buying situation including a specific brand or store. Following each buying situation were eight questions concerning four different risk producing dimensions. Based on the conceptual work of the Harvard group (Cox, 1967), each type of risk (social, etc.) was assumed to have an uncertainty component and a consequence component. Thus, two types of risk measures were required for each type of risk. The respondents were asked to indicate on a seven point scale their impressions of the likelihood of occurrence and how important it would be to avoid the consequences associated with each of the four risk producing dimensions, within the context of the given situation. Following is an example of the type of questions and scales used.

SITUATION: You are considering buying a new brand of electric knife at Gibson's Discount Store to give as a wedding present to a close friend.

How likely is it that others would think less of you if you were to give this product?

1 2 3 4 5 6 7  
Extremely Quite Somewhat Can't Somewhat Quite Extremely  
Unlikely Unlikely Unlikely Say Likely Likely Likely

How important is it for you to avoid causing others to think less of you?

1 2 3 4 5 6 7  
Extremely Quite Somewhat Can't Somewhat Quite Extremely  
Unimportant Unimportant Unimportant Say Important Important Important

The total response to any given risk producing dimension is considered to be the product of the numerical representations of the two questions. These measures of dimensional perceived risk will be considered as dependent variables for purposes of statistical analysis.

The four risk producing dimensions which were chosen for study were: social risk, financial risk, performance risk, and physical risk. Table 1 contains the operational definitions for each of these dimensions.

TABLE 1  
OPERATIONAL DEFINITIONS OF RISK DIMENSIONS

Risk Dimension	Operational Definition
PERFORMANCE RISK	How likely is it that this product is of poor quality? How important is it for you to avoid poor quality?
SOCIAL RISK	How likely is it that others would think less of you if you were to buy this product? How important is it for you to avoid causing others to think less of you?
PHYSICAL RISK	How likely is it that this product is hazardous to use? How important is it for you to avoid any such hazard?
FINANCIAL RISK	How likely is it that this product costs too much? How important is it for you to avoid paying too much?

### Statistical Procedure

The statistical analysis for the experiment was an analysis of variance with repeated measures. This type of analysis separates the respondents into groups according to the number of situations in the experiment. Since each respondent was presented only one situation and each received individual identification numbers, it was possible to separate the residual term into two distinct parts, between variations and within variations.

The structuring of F-statistics for an evaluation of brand or store effects was accomplished through division of the mean square due to brand or store effects by the mean square due to within variations. In a similar manner, the F-statistics for the evaluation of brand situational interaction or store situation interaction was accomplished through division of their respective mean squares by the mean square due to within variations. The F-statistics for the evaluation of situational effects was accomplished through division of the mean squares due to situational effects by the mean square due to between variations. A further discussion on repeated measures may be found in statistical principles in experimental design by B. J. Winer (1962).

### Results

The experiment concerns the purchase of an electric knife either for one's home use or as a wedding gift to a close friend. Each of forty-four individuals was given one of these two situations and responded to measures of performance risk, social risk, physical risk, and financial risk for two stores.

Table II contains the mean scores obtained for each dependent variable of the experiment and should be used in conjunction with the analysis of variance tables to indicate where differences exist and the direction of those differences. The higher the score, the greater is the level of perceived risk. Tables III through VI are the analysis of variance tables for each dependent variable or risk producing dimension. The percent contribution column was obtained utilizing an Omega Square statistic described by Hays (1964) and is intended to reflect the percentage of the total variation from the dependent variable which can be explained by each independent variable. This statistic has been calculated for only those effects which entered the analysis at a .05 or higher level of significance.

TABLE 2  
TABLE OF MEANS  
PERCEIVED RISK IN ELECTRIC KNIFE PURCHASE

Score	Situation	N	Social Risk	Financial Risk	Performance Risk	Physical Risk
Montgomery Ward	Home usage	22	11.27	27.04	17.32	20.91
Montgomery Ward	Wedding gift	22	15.64	19.14	19.73	18.23
Gibson's Discount	Home usage	22	9.68	26.86	22.64	22.09
Gibson's Discount	Wedding gift	22	19.18	16.23	21.45	21.04
Situation						
Home usage		44	10.48	26.95	19.98	21.50
Wedding gift		44	17.41	17.68	20.59	19.64
Store						
Montgomery Ward		44	13.45	23.09	18.52	19.57
Gibson's Discount		44	14.43	21.54	22.04	21.57

TABLE 3  
ANALYSIS OF VARIANCE FOR SOCIAL RISK:

Source	DF	Mean Square	F Value	Prob > F	Contribution (in Percent)
Situation	1	1057.10	9.23	.01	13.05
Between	42	114.57			
Store	1	21.01	.82	.63	
Situation x store	1	145.10	5.69	.02	1.6
Within	42	25.46			
Total	87	81.66			

TABLE 4  
ANALYSIS OF VARIANCE FOR FINANCIAL RISK:

Source	DF	Mean Square	F Value	Prob > F	Contribution (in Percent)
Situation	1	1891.64	9.41	.01	13.92
Between	42	200.94			
Store	1	52.54	1.46	.23	
Situation x store	1	40.91	1.13	.29	
Within	42	36.06			
Total	87	137.23			

TABLE 5  
ANALYSIS OF VARIANCE FOR PERFORMANCE RISK:

Source	DF	Mean Square	F Value	Prob > F	Contribution (in Percent)
Situation	1	8.28	.07	.78	
Between	42	115.13			
Store	1	273.01	5.44	.02	3.03
Situation x store	1	70.92	1.41	.24	
Within	42	50.19			
Total	87	83.86			

TABLE 6  
ANALYSIS OF VARIANCE FOR PHYSICAL RISK:

Source	DF	Mean Square	F Value	Prob > F	Contribution (in Percent)
Situation	1	76.41	.40	.53	
Between	42	189.20			
Store	1	88.00	1.96	.16	
Situation x store	1	14.73	.33	.58	
Within	42	44.81			
Total	87	115.03			

#### Social Risk

It is apparent that situational factors have entered at a highly significant level ( $P < .01$ ) as a means of explaining variations in responses to social risk. In this case, 13.05 of the total variation was explained by situational factors. The social risk involved in considering the purchase of an electric knife for home usage is significantly less than considering the same product as a wedding gift. It appears that the social risk involved in buying a product for oneself would be less than buying the same product as a gift for someone else. With respect to levels of social risk, there appeared to be no statistically significant store effect.

The interaction effects due to store-situation combinations significantly ( $P < .02$ ) explained 1.6% of the variation in social risk. This result would indicate that the experimental results attributed to the situational factor are slightly influenced by the store factor. In this specific case, it appears that the social risk involved in buying an electric knife for home usage is lower for Gibson's Discount store than for Montgomery Ward. However, the social risk asso-

ciated with buying an electric knife from Gibson's Discount store as a wedding present is significantly greater than from Montgomery Ward.

#### Financial Risk

Situational factors significantly ( $P < .01$ ) explained 13.92 of the variation in financial risk with financial risk being greater for the home usage situation than for the wedding gift situation. Here, the direction of mean differences is reversed from that obtained for social risk. This would imply that the respondents were more concerned about the price of a product which they intended to use themselves than when giving it as a wedding present.

In this particular case, by buying a more expensive model for a wedding present than they would have considered for themselves, could possibly be interpreted as a risk reduction method for handling perceived social risk.

There were no statistically significant store or interaction effects upon financial risk.

#### Performance and Physical Risk

With performance and physical risk, the only significant effect was related to store differences.

A significant ( $P < .02$ ) store effect on performance risk was indicated and explained 3.03% of the total variation for that dependent variable. Since the performance risk for an electric knife was less for Montgomery Ward than for Gibson's Discount, it would appear that these respondents felt that a better quality of electric knife could be found at a Montgomery Ward store than at a Gibson's Discount store.

No other main effects or interaction effects were found which significantly influenced the levels of perceived risk for either of these risk dimensions.

#### Discussion of Results

Four hypotheses concerning the independent effect of buying situations upon four dimensions of perceived risk were tested; two were statistically significant. When significant results were found, the amount of total variation explained by situation ranged from 13.05% to 13.92%. Store factors had a minor influence on performance risk. The magnitude of the variance explained supports Lutz and Kakkar's (1975) contention that one should not expect too much for situation in isolation. The results do, however, suggest that buying situations are related to "selected" intervening variables which may influence brand performance or purchase.

Theoretically, it appears that buying situations, along with individual differences and product characteristics, may play a part in determining perceived risk. This would imply that it would be appropriate to identify specific buying situations when experimental efforts are being made to measure perceived risk. This may eliminate a significant portion of unexplainable variations which has previously existed between individual respondents.

In a practical sense, it may be useful for advertisers to identify buying situations which tend to elicit high levels of perceived risk and either attempt to lessen these effects by illustrating the appropriateness of the product in the high risk situation, or simply choose to avoid advertisement formats which depict such high risk situations.

## References

- J. Arndt, "Word-of-Mouth Advertising and Informal Communication," In Donald F. Cox (ed.) Risk Taking and Information Handling in Consumer Behavior (Boston: Division of Research, Harvard Business School, 1967).
- Raymond A. Bauer, "Consumer Behavior as Risk Taking," Dynamic Marketing for a Changing World, R. S. Hancock (ed.) (Chicago: American Marketing Association, 1960), 389-398.
- Russell W. Belk, "An Exploratory Assessment of Situational Effects in Buyer Behavior," Journal of Marketing Research, XI(May, 1974), 156-163.
- Russell W. Belk, "The Objective Situation as a Determinant of Consumer Behavior," Advances in Consumer Research, Vol. II, M. J. Schlinger (ed.) (Chicago: 1975).
- Donald F. Cox, "Perceived Risk and Consumer Decision Making," Journal of Marketing Research, I(November, 1964), 32-39.
- Donald F. Cox, (ed.). Risk Taking and Information Handling in Consumer Behavior (Boston: Harvard University Press, 1967).
- S. M. Cunningham, "Perceived Risk as a Factor in Product-Oriented Word-of-Mouth Behavior: A First Step," In Peter Bennett (ed.) Proceedings American Marketing Association, 1965, 229-238.
- W. L. Hays, Statistics for Psychologists (Chicago: Holt, Rinehart, and Winston, 1964).
- Robert D. Hisrich, Ronald J. Dornoff, and Jerome B. Kernan, "Perceived Risk in Store Selection," Journal of Marketing Research, IX (November, 1972), 435-439.
- Richard J. Lutz and Pradeep Kakkar, "The Psychological Situation as a Determinant of Consumer Behavior," Advances in Consumer Research, Vol. II, M. Schlinger (ed.) (Chicago: 1975).
- Kenneth E. Miller, "A Situational Multi-Attribute Model," Advances--Consumer Research, Vol. II, M. Schlinger (ed.) (Chicago: 1975).
- Michael Perry and B. Curtis Hamm. "Canonical Analysis of Relations Between Socioeconomic Risk and Personal Influence in Purchase Decisions," Journal of Marketing Research, Vol. VI, (August, 1969), 351-354.
- Ted Roselius, "Consumer Rankings of Risk Reduction Methods," Journal of Marketing, XXXV (January, 1971), 56-61.
- Ivan Ross, Perceived Risk and Consumer Behavior: A Critical Review, Working Paper Number 21. (Minneapolis: University of Minnesota, 1974).
- R. G. Sandell, "Effect of Attitudinal and Situational Factors on Reported Choice Behavior," Journal of Marketing Research, V, 1968, 405-408.
- H. E. Spence, J. F. Engel, and R. D. Blackwell, "Perceived Risk in Mail-Order and Retail Store Buying," Journal of Marketing Research, VII (August, 1970), 364-369.
- Bernard J. Winer, Statistical Principles in Experimental Design. (New York: McGraw-Hill, 1962), 298-318.
- W. G. Zikmund and J. E. Scott, "A Multivariate Analysis of Perceived Risk, Self-Confidence and Information Sources," Advances in Consumer Research, Volume I. Proceedings of the Association for Consumer Research. S. Ward and P. Wright (ed.) (Chicago: 1973), 406-416.



## DEMAND BIAS IN THE ASSESSMENT OF SITUATIONAL EFFECTS ON BUYER BEHAVIOR

Peter H. Reingen, Iona College

### Abstract

The possibility that the conventional assessment of situational effects on buyer behavior introduces demand bias was investigated. Both demand awareness and subject role set were manipulated in four conditions. The results substantiated that possibility.

### Introduction

Only recently has any systematic research been undertaken to assess the impact of situational variables on buyer behavior (Belk, 1974a,b; Belk, 1975; Lutz and Kakkar, 1975). Based on the premise that consumer behavior is a function of the interaction between the individual and the situation, this research provides evidence that substantially more behavioral variance can be accounted for by explicitly introducing situational variables into the research design rather than by ignoring them. Thus, this line of research offers much promise.

In the standard investigation of situational effects (e.g., Belk, 1974a), the subjects respond to a modified behavioral differential questionnaire in which the likelihood is indicated that they would choose each of a number of responses (products) to a stimulus configuration (situation). The result is a treatments (situations) by treatments (products) by subjects design which allows for the estimation of the relative contribution of the main and two-way interaction effects to the variance in the dependent variable. The patterns of variance are then examined with the use of multimode factor analysis which renders an interpretation of the types of products preferred by certain types of buyers in various types of situations.

Results appear to confirm the dominance of interactions over main source effects in contribution to variance. For example, the Situations by Products interaction has been found to be an important determinant of choice behavior for snack and meat products inventories (Belk, 1974a). And this is the effect commonly implied when one refers to situational influences in buyer behavior. However, Lutz and Kakkar (1975) report an interaction between Situations and Products accounting for considerably less variance than found by Belk (1974a). While there are alternative explanations for this discrepancy (e.g., the two studies employed different methods), Lutz and Kakkar's (1975) suggestion that the difference in results may have been due to demand characteristics that were operating in Belk's (1974a) experiment deserves careful investigation to assure future progress along this promising line of research.

The concept of demand characteristics has Lewinian-Festalk connotations (Orne, 1970, p. 226), and it concerns socio-psychological aspects of the experiment which may be used by the subjects as cues to penetrate the ambiguousness or deceptiveness of instructions and to discern the experimenter's expectancies regarding their behavior. Demand characteristics, when allowed to operate unnoticed and uncontrolled by the experimenter, can contaminate an otherwise well-designed experiment, and they can pose strong threats to both internal and external validity. For example, Page (1974) and Page and Scheidt (1971) provide evidence that de-

mand characteristics confounded experiments on the classical conditioning of attitudes (e.g., A.W. Staats and C.K. Staats, 1958) and the so-called weapons effect (Berkowitz and LePage, 1967). Consumer researchers as well have recently addressed themselves to the problem of demand characteristics (e.g., Sawyer, 1975a,b).

The argument that Belk's (1974a) results may be vulnerable to a demand effect interpretation is founded on the modified behavioral differential questionnaires the study employed. In those questionnaires, the situations vary while the product choices remain constant. When a within-subjects design is employed, this may provide the subjects with the cue that they are "supposed to" shift in their buying choices across the situations. Based on Orne's (1962) suggestion that most subjects are motivated to try to confirm what they believe to be the experimental hypothesis, the result may be an artificially inflated Situations by Products interaction effect. However, while many subjects appear to adopt such a "compliant" role (e.g., Levy, 1967; Sawyer, 1975a), other roles with different effects on the dependent variable have been identified. The "negative" subject role occurs when a subject tries to disconfirm a suspected experimental hypothesis (Allen, 1967; Berkowitz, 1971), while the "faithful" subject attempts to continue in a role that emphasizes adherence to instructions, despite suspicion, for the sake of the validity of the experiment (Fillenbaum, 1966; Golding and Lichtenstein, 1970). Thus, the specific effects of demand awareness depend on adopted subject roles. Demand awareness may cause false positive findings ("compliant" subject role), false negative findings ("negative" subject role), or it may have no important performance effects ("faithful" subject role).

### Testing for Demand Artifacts

These various considerations led to the following experiment in which both demand awareness and subject role set were manipulated in four conditions. In the first condition, subjects completed a product inventory taken from the Belk (1974a) study and they responded to a brief post-experimental questionnaire which was designed to isolate subjects who were aware of the experimenter's expectation that their product preferences would vary across situations (Control Condition).<sup>1</sup> The subjects in the second condition were informed about the experimenter's expectation before their completion of the questionnaire (Aware Condition). In the third condition, subjects were instructed to confirm what they thought the experimenter's expectation was (Confirm Condition), while in the fourth condition they were asked to disconfirm it (Disconfirm Condition). The subjects in the third and fourth conditions also completed the post-experimental questionnaire administered to subjects in the first condition.

<sup>1</sup>This is not to suggest that Belk (1974a) specifically tested such a hypothesis. Rather, based on Lutz and Kakkar's (1975) suggestion that demand awareness led to the greater Situations by Products interaction effects found by Belk (1974a), this was construed to be the experimenter's hypothesis.

Since the preponderance of evidence suggests a "compliant" subject role, the variance accounted for by the Situations by Products interaction should be greater in the Aware Condition than in the Control Condition. Similarly, under the assumption that a substantial number of subjects are demand-aware, this interaction should explain more variance in the Confirm Condition than in the Control Condition. Alternatively, in the Disconfirm Condition, the importance of this interaction should be less than in the Control Condition. If these results should not materialize, the Control Condition's outcome would likely be vulnerable to a demand bias interpretation.

#### Method

The subjects, 101 undergraduates from introductory business courses at Iona College, were assigned to four conditions: Control Condition (24 subjects), Confirm Condition (28 subjects), Disconfirm Condition (25 subjects), and Aware Condition (24 subjects). Each subject was administered the Belk (1974a) snack products questionnaire which consists of ten different snack consumption situations. The subjects were asked to respond to a ten-item (snack products) behavioral differential inventory in each situation. The subjects in the Control Condition received the following general instructions:

"On the following pages you will be asked to check how likely it would be that you would choose snacks in a series of circumstances. For example, friends have invited you to a picnic and asked you to bring some snacks along. How likely would it be that you would choose each of the following snacks in these circumstances: potato chips, popcorn, cookies, fresh fruit, sandwiches, pastries, ice cream, cheese, assorted nuts, and crackers. In the questionnaire, there are ten different circumstances and for each one of them you will be asked how likely it would be that you would choose each of these snacks."

The subjects in the Confirm Condition, Disconfirm Condition, and Aware Condition received these general instructions as well as the following condition-specific instructions:

"As you can imagine I have a certain expectation as to how the results will turn out. Needless to say, I can't tell you about my expectation beforehand, that is, the way I think you will fill out the questionnaire; but while you are filling out the questionnaire you may get some idea as to what you think my expectation is. If you should get an idea, fill out the questionnaire according to what you think my expectation is. If not, fill it out anyway." (Confirm Condition.)

"As you can imagine I have a certain expectation as to how the results will turn out. Needless to say, I can't tell you about the expectation beforehand, that is, the way I think you will fill out the questionnaire; but while you are filling out the questionnaire you may get some idea as to what you think my expectation is. If you should get an idea, fill out the questionnaire opposite to what you think my expectation is. If not, fill out the questionnaire anyway." (Disconfirm Condition.)

"As you can imagine I have a certain expectation as to how the results will turn out. I may as well tell you about my expectation beforehand. I expect that the likelihood of your choosing a particular snack is dependent upon the circumstances. For example, when friends have invited you to a picnic, you may be more likely to choose cheese as a snack to bring along than

ice cream. Under different circumstances, however, you may be more likely to choose ice cream as a snack than cheese." (Aware Condition.)

After the subjects in the Control Condition, Confirm Condition, and Disconfirm Condition finished the snack products instrument, they were instructed to complete the following questionnaire:

1. When I presented you with the questionnaire, suppose that I expected that the likelihood of your choosing a particular snack is dependent on the circumstances. While you were filling out the questionnaire, did you expect that this was my approximate hypothesis?  
Yes \_\_\_ No \_\_\_

2. If yes, how certain were you of this at the time you were filling out the questionnaire?

Guessing     :     :     :     :     :     :     Certain  
          1      2      3      4      5      6      7

3. As you can see, one thing I am studying concerns your suspicions about the questionnaire. Make any other comments here about your thoughts, reactions, and suspicions about the questionnaire.

A subject was considered demand-aware if he responded affirmatively to Question 1 and received a score of at least five for Question 2. A score of at least five should be indicative of an acceptable level of certainty in the subject's response to Question 1.

#### Results

Based on the data secured by the post-experimental questionnaire and the previously stated criteria, 17 out of 23 subjects (73%) in the Control Condition, 10 of 25 subjects (40%) in the Confirm Condition, and 12 of 25 subjects (48%) in the Disconfirm Condition were classified as demand-aware ( $\chi^2 = 5.986$ ;  $df = 2$ ;  $p > .05$ ).<sup>2</sup> This suggests that demand-awareness may contaminate results obtained by the conventional assessment of situational effects on buyer behavior. Consequently, further analysis was performed to determine the specific effects of demand-awareness.

For each condition separately, the subjects' choice data were analyzed with the intent to construct estimates of the relative contribution of sources -- as derived from expected mean square formulas specified by a three-way random effects model (Gleser, Cronbach, and Rajaratam, 1965) -- to variance in the dependent measure.<sup>3</sup>

The results in Table 1 show that the relative contributions of individual sources to variance in product preferences varied markedly across the four conditions. Generally speaking, this indicates that different role sets and levels of demand awareness may contaminate overall results. With respect to the source effect of primary interest, namely the Situations by Products interaction effect, its relative importance differed across the conditions. More specifically, the Situations by Products interaction contributed 1.49% (Disconfirm Condition), 11.49% (Aware Condition), 14.20%

<sup>2</sup>Four incomplete questionnaires (one, three, and one for the Control Condition, Confirm Condition, and Aware Condition, respectively) were discarded. Responses to the third item of the post-experimental questionnaire such as "I figured that you wanted to know what snack I would pick under a certain set of circumstances since the types of snack remained the same and only the circumstances changed" were not uncommon.

TABLE 1

## RELATIVE CONTRIBUTIONS OF SOURCES TO VARIANCE IN PRODUCT PREFERENCES

Source	Control		Confirm		Disconfirm		Aware	
	a MS	b %	MS	%	MS	%	MS	%
Persons	9.49	.48	9.37	.52	27.44	5.75	44.02	14.48
Situations	17.75	1.13	6.96	0.00	1.98	0.00	9.72	.16
Products	35.72	3.30	24.24	1.43	9.46	0.00	26.56	2.30
Situations by Products	10.59	16.34	9.46	14.20	3.12	1.49	8.06	11.49
Situations by Persons	1.97	2.90	1.66	1.77	3.74	5.75	2.00	2.81
Products by Persons	7.59	25.52	7.71	27.96	8.30	20.84	6.14	18.92
Residual	1.25	50.32	1.25	54.11	2.00	66.16	1.28	49.84

<sup>a</sup>MS = Mean Square

<sup>b</sup>Percent Contribution

(Confirm Condition), and 16.34% (Control Condition) to the variance in the dependent measure. These findings imply that the Control Condition's outcome is very likely not vulnerable to a "negative" subject interpretation but more vulnerable to a "compliant" one.

#### Discussion

The considerable proportions of subjects identified here as demand-aware leave previous results (e.g., Belk, 1974a) open to a demand bias interpretation. One could argue that post-experimental questionnaires -- especially when lengthy and detailed -- have the undesirable effect of "suggesting" awareness to subjects. However, in light of the very brief questionnaire that was employed in this study, such a possibility appears to be weak. Moreover, as the discussion that follows indicates, the negligible Situations by Products interaction effect obtained in the Disconfirm Condition does not support the notion that the post-experimental questionnaire "suggested" awareness to many of those subjects who were classified as demand-aware.

As far as the specific effects of demand awareness are concerned, the substantially lower variance accounted for by the Situations by Products interaction in the Disconfirm Condition strongly suggests that the Control Condition's result was not suppressed by a dominance of subjects who adopted a "negative" role. Furthermore, this small interaction effect indicates that the subjects in the Disconfirm Condition classified as demand-aware apparently became so soon after they had started

to complete the questionnaire and could readily manipulate their answers in accordance with their perception of the experimenter's expectation. There is no compelling reason to believe that the demand-aware subjects in the Confirm Condition differed substantially from the demand-aware subjects in the Disconfirm Condition in regard to overall data manipulation ability and the experimenter expectation recognition pattern. Consequently, if a demand bias interpretation due to "compliant" subjects is to be ruled out, the variance accounted for by the Situations by Products interaction should have been much greater in the Confirm Condition than in the Control Condition. Quite to the contrary, however, this interaction accounted for slightly more variance in the Control Condition than in the Confirm Condition. This result is plausible in light of the greater proportion of demand-aware subjects in the former condition and the findings which suggest that more subjects adopt a "compliant" role (e.g., Sawyer, 1975a), but it is not necessarily consistent with the outcome of the Aware Condition. In the Aware Condition, the variance accounted for by the Situations by Products interaction was less than in the Control Condition. In accord with Brehm's (1966) reactance theory, some subjects might regard the explicit introduction of the experimenter's expectation as an unwelcome pressure on them. In resisting this pressure, they may then show relatively little of the behavior the experimenter wants.

In sum, this research shows that the conventional assessment of situational effects in buyer behavior may introduce demand characteristics with confounding effects. As indicated by the great proportions of subjects who were classified as demand-aware, the within-subjects design -- where a number of experimental treatments (i.e., situations) is administered to the same subject -- commonly employed in this assessment is particularly demand-prone. This calls for a greater reliance on between-subjects designs in future research. Alternatively, the number of treatments to which an individual subject is exposed should be reduced. The greater the number of treatments in a within-subjects design, the more likely is a subject to successfully estimate the experimenter's expectations. While the results render very little evidence in favor of a "negative" subject role and more overall evidence in support of a "compliant" subject role interpretation, this study does not provide conclusive evidence as to the specific effects of operating demand characteristics in the conventional assessment of situational vari-

<sup>3</sup>Since the situations and products in Belk's (1974a) inventory are not random samples of all possible situations and products, a mixed-effects model may appear to be more appropriate. However, in contrast to the random-effects model, with the mixed-effects model one can solve for all the components of variance only if one assumes the triple interaction to be zero. This triple interaction should be meaningful, and consequently it was believed to be less hazardous to assume a random-effects model. At any rate, as Endler and Hunt (1966) point out, the components of variance are comparable for random-effects and mixed-effects models. Thus, even if a mixed-effects model would have been assumed, the interpretation of the results across the conditions would have remained unchanged.

ables. Regardless of specific effects, the researcher's best assumption is probably that data from suspicious subjects are untrustworthy.

#### References

- Allen, V.L. "The Effect of Knowledge of Deception on Conformity," Journal of Social Psychology, 69 (1966), 101-106.
- Belk, R.W. "An Exploratory Assessment of Situational Effects in Buyer Behavior," Journal of Marketing Research, 11 (May 1974), 156-163(a).
- Belk, R.W. "Application and Analysis of the Behavioral Differential Inventory for Assessing Situational Effects in Buyer Behavior," in S. Ward and P. Wright, Advances in Consumer Research: Volume I. Urbana, Ill.: Association for Consumer Research, 1974, 360-370(b).
- Berkowitz, L. and A. LePage. "Weapons as Aggression-Eliciting Stimuli," Journal of Personality and Social Psychology, 7 (1967), 202-207.
- Berkowitz, L. "The Weapons Effect, Demand Characteristics, and the Myth of the Compliant Subject," Journal of Personality and Social Psychology, 20 (1971), 332-338.
- Brehm, J.A. A Theory of Psychological Reactance. New York: Academic Press, 1966.
- Endler, N.S. and J.M. Hunt. "Sources of Behavioral Variance as Measured by the S-R Inventory of Anxiousness," Psychological Bulletin, 65 (June 1966), 336-346.
- Fillenbaum, S. "Prior Deception and Subsequent Experimental Performance: The Faithful Subject," Journal of Personality and Social Psychology, 4 (1966), 532-537.
- Glesner, G.C., L.J. Cronbach, and N. Rajaratnam. "Generalizability of Scores Influenced by Multiple Sources of Variance," Psychometrika, 30 (December 1965), 395-418.
- Golding, S.L. and E. Lichenstein. "Confession of Awareness and Prior Knowledge of Deception as a Function of Interview Set and Approval of Motivation," Journal of Personality and Social Psychology, 14 (1970), 213-223.
- Levy, L. "Awareness, Learning and the Beneficent Subject as Expert Witness," Journal of Personality and Social Psychology, 6 (1967), 363-370.
- Lutz, J.R. and P. Kakkar. "The Psychological Situation as a Determinant of Consumer Behavior," in M.J. Schlinger, Advances in Consumer Research: Volume II. Chicago, Ill.: Association for Consumer Research, 1975, 439-454.
- Orne, M. "On the Social Psychology of the Psychological Experiment: With Specific Reference to Demand Characteristics," American Psychologist, 17 (November 1962), 776-783.
- Orne, M. "Hypnosis, Motivation, and the Ecological Validity of the Psychological Experiment," in W.J. Arnold and M.M. Page, Nebraska Symposium on Motivation. Lincoln, Nebraska: University of Nebraska Press, 1970.
- Page, M.M. and R.J. Scheidt. "The Elusive Weapons Effect: Demand Awareness, Evaluation Apprehension, and Slightly Sophisticated Subjects," Journal of Personality and Social Psychology, 30 (1974), 468-476.
- Sawyer, A.G. "Detecting Demand Characteristics in Laboratory Experiments: The Case of Repetition-Affect Research," in M.J. Schlinger, Advances in Consumer Research: Volume II. Chicago, Ill.: Association for Consumer Research, 1975, 713-724(a).
- Sawyer, A.G. "Demand Artifacts in Consumer Research," Journal of Consumer Research, 1 (March 1975), 20-30(b).
- Staats, A.W. and C.K. Staats. "Attitudes Established by Classical Conditioning," Journal of Abnormal and Social Psychology, 57 (1958), 37-40.

## DISSONANCE RESOLUTION BY GRADE SCHOOL CONSUMERS

David B. Kyner, DKG Advertising  
Jacob Jacoby and Robert W. Chestnut, Purdue University

### Abstract

Children evaluated two candy bar brands, made six purchase decisions involving these brands, then re-evaluated the same two brands. Significant ( $p < .01$ ) pre-to-post attitude shifts occurred as predicted by dissonance theory: shifts in predicted directions were greater when the brands were initially considered to be equally attractive than when they were considered unequally attractive.

### Introduction

Children, as a focus for consumer behavior research, have been relatively neglected (cf., McNeal, 1969; Jacoby and Kyner, 1973). While research in related areas such as advertising impact has helped to provide data on consumer-related experiences during childhood (Ward, Robertson and Wackman, 1971; Robertson and Rossiter, 1974), attention has only recently been given to the cognitive processes by which a child arrives at and possibly re-evaluates a consumer decision (Calder, Robertson and Rossiter, 1975).

Information on such processes would be useful from several standpoints. First, children represent an important market segment. They spend nearly 500 million dollars a year in their own right (cf., Jacoby and Kyner, 1973, p. 4) and exert considerable influence over expenditures made by other family members (Coulson, 1966; Deering and Jacoby, 1971; McNeal, 1969; Munn, 1958; Wells and LoSciuto, 1966). Knowing how a child will evaluate and influence the purchase of a given product should significantly improve the marketer's ability to present the product in a favorable light. Second, governmental interest in consumer safety and education often center on the child as consumer. Consider, for example, the FTC's increasing concern over advertising directed at children. Yet basic information regarding how the child perceives products and acts within his role as a consumer is still lacking, and such knowledge is vital for arriving at appropriate public policy decisions. Finally, from a longitudinal perspective, research on consumer behavior in children can provide insight into how adult buying behaviors are acquired. Consumer behaviorists generally assume that certain aspects of consumer behavior are acquired sometime during childhood, yet they typically fail to show how these behaviors are learned or even if they exist as important elements in the child's behavior as a consumer.

The present investigation focuses upon one cognitive process that has been shown to have a significant effect on adult consumer behavior. The process is known as post-decision dissonance reduction and, in the consumer context, involves a drive toward re-evaluating brand preferences immediately after purchase. Of all the empirical issues and research stimulated by Festinger's dissonance theory, reviewers (e.g., Insko, 1967, p. 283; Kiesler, Collins and Miller, 1969, p. 205) consistently point to post-decision dissonance reduction as one of the more stable and interesting research issues. Although an increasing number of studies have appeared on the subject, a literature review reveals no attempt to extend research on this issue to children.

Early cognitive dissonance investigations (Brehm, 1956) demonstrated a tendency to re-evaluate alternatives immediately following a choice situation: the chosen alternative increased in desirability while the rejected alternative became less desirable. In terms of dissonance theory, it was assumed that, after making a choice, a person would attempt to justify his decision by emphasizing the positive aspects of the chosen alternative and the negative aspects of the rejected alternative. This in turn led to a change in judged desirability. Dissonance theory predicts that the more nearly equal the attractiveness of two alternatives, the more dramatic the post behavioral re-evaluation of attitudes toward these alternatives -- in a favorable direction for the selected alternative and in an unfavorable direction for the rejected alternative. Another prediction is that, as the cognitive overlap (or actual similarity) of the two alternatives decreased, the dissonance and resultant re-evaluation would increase (Brehm and Cohen, 1962, p. 303). The case of maximal dissonance, therefore, would be a choice between two equally desirable alternatives (i.e., both highly attractive) in which the attributes of one alternative were different from those of the second alternative.

Empirical support for these hypotheses was adduced by Brehm and Cohen (1959, 1962) and Brock (1963). These first tests of the theory often utilized children. Brock, for example, tested 141 children from three to twelve years of age. In one condition, children were to choose between similar objects (two crackers or two toys), while in the other condition, they were to choose between dissimilar objects (a cracker and a toy). In the usual comparison of pre vs. post-ratings of desirability, children choosing between dissimilar objects reflected a greater change in attitudes and, thus, were assumed to have experienced more dissonance.

Brock's experiment would seem to indicate that post-decision dissonance reduction via a change in perceived desirability plays an important role in a child's decision-making and that it manifests itself at a very early age. Unfortunately, this study has never been replicated in a consumer choice-situation.

Consumer studies which did replicate Brehm's findings were all concerned solely with adult behavior. Anderson, Taylor and Holloway (1966), varying both the number of alternatives and relative attractiveness, showed the predicted changes in desirability on a wide variety of products. LoSciuto and Perloff (1967) provided additional supportive evidence through the use of a ranking task which required people to rank in order the desirability of nine record albums. A high dissonance condition was created by arranging a choice between the third and fourth ranked alternatives, while a low dissonance condition involved a choice between the third and eighth ranked alternatives. As predicted, a greater shift in desirability ratings was noted in the high dissonance condition. Sheth (1970) replicated this finding on a sample containing both housewives and male graduate students.

Despite criticisms that the LoSciuto and Perloff design incorporates a biasing "regression effect," (Oshikawa, 1971), rejoinders (Sheth, 1971; Cummings and Venkatesan, 1974) suggest that the early studies

by Brehm and his associates appear to have carried over quite well to the consumer context when adult consumers are considered. Re-examination of data collected earlier for another purpose (cf., Jacoby and Kyner, 1973), permits a test of the following hypothesis: as the desirability of two purchase alternatives becomes more nearly equal, dissonance will increase for children engaged in a consumer decision making task.

## Method

### Subjects

The subjects were 80 six-to-nine year old children, 32 males and 48 females.<sup>1</sup> Previous studies (Vogl, 1964; Marshall and Magruder, 1960) had indicated that children of this age were both selective in the brands they chose and knowledgeable in the use of money. This finding was confirmed in pilot tests with subjects similar to those used in the present study (Kyner, 1971), as well as in discussions with the subjects employed in the present investigation.

### Design and Procedure

The experiment was conducted in three segments: a pre-rating of brand preference, a set of six purchase trials, and a post-rating of preference. Each child was tested independently.

Children were first exposed to two different brands of candy bars. These brands had similar appearance and were identified only by the labels L and S -- letters found to elicit relatively neutral affect and low candy associations in the pilot studies (cf., Kyner, 1971). In actuality, brands L and S were the same candy bar. The pilot studies had shown, however, that children still managed to perceive differences and that these differences were randomly distributed. After presenting the supposedly different brands, children were asked to taste-test the candy bars and rate them via two scales: a seven-point Smiling Faces Scale (SFS; cf., Wells, 1965) and a ten-point Method of Constant Sum (MCS). Each child's most preferred brand (MPB) and second preferred brand (SPB) was identified via these ratings.

Immediately following this initial rating, the children were engaged in a series of six purchase trials. These trials were administered at four minute intervals during a videotaped cartoon show. For the first five trials the child was confronted with a short, mildly persuasive videotaped advertisement for his MPB after four minutes of cartoons had elapsed. The child was then given a dime and asked to use it in making a purchase decision between the two brands. The context of the first five trials, therefore, tended to reinforce brand loyalty to the MPB. This was successful in that 89% of the 400 choices (80 subjects x 5 trials) were for the MPB. Finally, a sixth trial was included that followed a longer advertisement which contained a strong appeal against the MPB and advocated purchase of the SPB.

The third stage of the experiment consisted of administering a number of measures, including both the SFS and MCS. These procedures are more fully described in Jacoby and Kyner (1973).

<sup>1</sup> Four subjects rated brands L and S as equal in the initial taste test and, thus, were forced into making a MPB selection. Since this represents a non-volitional choice (i.e., outside the realm of dissonance theory predictions), data from these subjects were excluded from the analysis.

## Results

Subjects were divided into two groups based upon their initial brand ratings. The "equal attractiveness" group consisted of all subjects who perceived little or no difference between brands in the initial taste-test. Operationally, these were subjects who rated the two brands as 0 or 1 point different on the SFS and 0 to 2 points different on the MCS. The "unequal attractiveness" group was composed of the remaining subjects, all of whom had differentiated between the brands. Four subjects had shown equal preference for the two brands on both the SFS and the MCS and, thus, were excluded from the analysis. For ease of comparison, the same type of analyses as employed by Brehm and Cohen (1959) were employed here.

Table 1 shows the results of a t-test between pre- vs. post-ratings on the seven-point SFS for both groups. As can be seen by the index of change, subjects perceiving no initial brand differences manifested the predicted shifts in preference. They increased their liking for the MPB and decreased their liking for the SPB. This change was significant at the .001 level for the equal attractiveness group and nonsignificant for the unequal attractiveness group. A t-test between the two groups showed a difference significant at the .01 level.<sup>2</sup>

Table 2 shows the same results for the ten-point MCS measure. Here again, t-tests for the equal attractiveness group were significant at the .001 level, while being nonsignificant for the unequal attractiveness group. A t-test between groups showed a significant difference at the .05 level.<sup>2</sup>

TABLE 1

Summary of t-tests for the two conditions of attractiveness as measured by the SFS

Attrac- tiveness Condition	N	Pre-mean rating	Post-mean rating	Net Change	df	t	p
Equal							
MPB	58	2.03	1.56	+ .47	57	5.17	.001
SPB	58	2.67	3.29	- .62	57	-5.08	.001
Both Brands	58			+1.09	74	2.52	.01
Unequal							
MPB	18	1.67	1.55	+ .12	17	.40	N.S.
SPB	18	4.05	4.16	- .11	17	-.32	N.S.
Both Brands	18			+ .23			N.S.

<sup>2</sup> A one-tailed t statistic for non-correlated data.

TABLE 2

Summary of t-tests for the two conditions of attractiveness as measured by the MCS

Attractiveness Condition	N	Pre-mean rating	Post-mean rating	Net Change	df	t	p
Equal							
MPB	57	5.95	6.42	+ .47	56	3.53	.001
SPB	57	4.05	3.58	- .47	56	-3.53	.001
Both Brands	57			+ .94	74	2.05	.05
Unequal							
MPB	19	7.63	7.53	- .10	18	- .38	N.S.
SPB	19	2.37	2.47	+ .10	18	.38	N.S.
Both Brands	19			- .20			N.S.

### Discussion

The hypothesis that, as the desirability of the available purchase alternatives becomes increasingly similar, post-decision dissonance reduction tendencies will increase was supported by the finding of this investigation. This was manifested by subjects rating their MPB higher and their SPB lower than they did initially. Such attitude changes occurred to a significantly greater degree when the most and second preferred brand were equally attractive than when they were unequally attractive.

Dissonance theory would reason that the positive attributes of the rejected alternative were greater in number for the equal attractiveness group. These positive attributes would generate greater dissonance for the child, as he was confronted with the fact of actually rejecting the SPB in the six purchase trials. To resolve this dissonance, the child apparently engaged in a re-evaluation of brand preferences.

Alternative explanations for this re-evaluation seem unlikely. Statistical regression, for example, would be inappropriate in explaining the systematic polarization of post-task ratings (cf., Campbell and Stanley, 1963, p. 10). A learning interpretation of the data, although of some predictive value, would not adequately account for the increasing negative evaluation of the SPB. Finally, the reliability of results from the two different rating procedures and the fact that post-task ratings did differ substantially from the end-points of the scales argue against potential ceiling effects having affected the results.

Given a dissonance orientation towards the data, one interesting finding was that, despite a great deal of cognitive overlap (i.e., the candy bars were the same in terms of physical appearance and actual composition), post-decision dissonance reduction still occurred. This would seem to conflict with Brehm's statement that "when the two alternatives have the same attributes -- that is, they are identical -- nothing is given up or avoided by choosing one over the other (1962, p. 37)." Dissonance should not occur in the case of identical alternatives.

It can be contended, however, that the brands in question were not perceived to be identical by the

subjects. They were presented under two different brand names (L & S) which seemed to generate a psychological reality of difference for the children. Not a single child ever questioned the fact that L and S were different, either during the experiment or in post-experimental interview and debriefing sessions. Taken in conjunction with the fact that post-decision dissonance can be assumed to have occurred, this would suggest that overlap must be defined in terms of "psychological" rather than "objective" reality.

In sum, this study has made two contributions to the previous literature on post-decision dissonance reduction. First, it has generalized the findings of Brehm and his associates to the consumer behavior of children by supporting the hypothesis that the closer the two alternatives are in attractiveness, the greater the subsequent attitude change. Second, it has demonstrated (at least within the context of a child's consumer behavior) that actual product similarity may have little effect upon the arousal of dissonance.

Both findings are important for future research involving dissonance reduction in the purchasing behavior of children. Above all, they suggest that investigators concentrate on the psychological reality of a brand within a product class. What, for instance, is the relationship between brand loyalty and a child's perception of similarity between brands? Can brand loyalty be reduced by actually informing the child of a lack of objective difference in quality? Because of their potential to provide useful insight into the development of various consumer behavior phenomena in adults (e.g., brand loyalty), a research effort designed to examine the implications of cognitive dissonance in the consumer behavior of children seems appropriate.

### References

- Lee K. Anderson, James R. Taylor and Robert J. Holloway, "The Consumer and His Alternatives: An Experimental Approach," Journal of Marketing Research, 3(February, 1966), 62-67.
- Jack W. Brehm, "Post-Decision Changes in the Desirability of Alternatives," Journal of Abnormal and Social Psychology, 52(May, 1956), 384-389.
- Jack W. Brehm and Arthur R. Cohen, "Reevaluation of Choice Alternatives as a Function of Their Number and Qualitative Similarity," Journal of Abnormal and Social Psychology, 58(May, 1959), 373-378.
- Jack W. Brehm and Arthur R. Cohen, Exploration in Cognitive Dissonance (New York: Wiley, 1962).
- Timothy C. Brock, "Effects of Prior Dishonesty on Postdecision Dissonance," Journal of Abnormal and Social Psychology, 66(April, 1963), 325-331.
- Bobby J. Calder, Thomas S. Robertson and John R. Rossiter, "Children's Consumer Information Processing," Communication Research, 2(July, 1975), 307-316.
- Donald T. Campbell and Julian C. Stanley, Experimental and Quasi-Experimental Designs for Research (Chicago: Rand McNally, 1963).
- John S. Coulson, "Buying Decisions Within the Family and the Consumer-Brand Relationship," in Joseph Newman (Ed.), On Knowing the Consumer (New York: Wiley, 1966).

William H. Cummings and M. Venkatesan, "Cognitive Dissonance and Consumer Behavior: A Critical Review," The University of Iowa Bureau of Business and Economic Research, Working Paper Series No. 74-1, (January, 1974).

Barbara J. Deering and Jacob Jacoby, "The Effects of Alternative Relationships and Relative Resources on Consumer Decisions Between Mother and Child," Proceedings. Second Annual Convention, The Association for Consumer Research, (1971),135-142.

Chester A. Insko, Theories of Attitude Change (New York: Appleton-Century-Crofts, 1967).

Jacob Jacoby and David B. Kyner, "Brand Loyalty vs. Repeat Purchasing Behavior," Journal of Marketing Research, 10(February, 1973),1-9.

Charles A. Kiesler, Barry E. Collins and Norman Miller, Attitude Change (New York: Wiley, 1969).

David B. Kyner, "Testing Jacoby's Conceptualization of Brand Loyalty," unpublished Ph.D. dissertation, Purdue University, 1971.

Leonard A. LoSciuto and Robert Perloff, "Influence of Product Preference on Dissonance Reduction," Journal of Marketing Research, 4(August, 1967),286-290.

Helen R. Marshall and Lucille Magruder, "Relations Between Parent Money Education Practices and Children's Knowledge and Use of Money," Child Development, 31 (1960),253-284.

James U. McNeal, Dimensions of Consumer Behavior (New York: Appleton-Century-Crofts, 1969).

Mark Munn, "The Effect of Parental Buying Habits of Children Exposed to Children's Television Programs," Journal of Broadcasting, 2(1958),253-258.

Sadaomi Oshikawa, "Dissonance Reduction or Artifact?," Journal of Marketing Research, 8(November, 1971),514-515.

Thomas S. Robertson and John R. Rossiter, "Children and Commercial Persuasion: An Attribution Theory Analysis," Journal of Consumer Research, 1(1974), 13-20.

Jagdish N. Sheth, "Are There Differences in Dissonance Reduction Behavior Between Students and Housewives?," Journal of Marketing Research, 7(May, 1970), 243-245.

Jagdish N. Sheth, "Dissonance Reduction or Artifact? A Reply," Journal of Marketing Research, 8(November, 1971),516-517.

A. J. Vogl, "The Changing Face of the Children's Market," Sales Management, 93(December 18, 1964), 35-38.

Scott Ward, Thomas S. Robertson and Daniel Wackman, "Children's Attention to Television Advertising," Proceedings. Second Annual Conference, The Association for Consumer Research, (1971),143-156.

William D. Wells, "Communicating with Children," Journal of Advertising Research, 5(June, 1965),2-14.

William D. Wells and Leonard A. LoSciuto, "Direct Observation of Purchasing Behavior," Journal of Marketing Research, 3(August, 1966),227-233.



## IDEAL POINT MODELS OF PREFERENCE

Joel Huber, Purdue University

### Abstract

Ideal point models possess an intuitive appeal which has contributed to their widespread acceptance by those modeling consumer behavior. This paper presents several technical suggestions for improving the models, but argues that conceptual problems severely limit their use, particularly in new product applications.

### Introduction

In an ideal point model, preference for a stimulus is represented as being inversely related to its distance from an ideal point within a stimulus space. This space is defined by the attributes of the stimulus set so that differences in preferences can be seen as the result of differing responses to those attributes. In a marketing context, the ideal point model provides an appealing geometric metaphor which can be used for defining new products, repositioning old products, and determining "benefit" segments who desire similar attributes in a product.

In spite of its intuitive appeal, the concept has questionable validity. Predictive studies by Jones and Pessemer (1974), dealing with text books; Braun and Srinivasan (1973), on grooming products; and Green and Jain (1972), on dessert items, have produced predictions better than chance, but not much. This paper considers two reasons for the disappointing performance of the model. First, there are technical problems with the routines used to estimate the ideal points. Second, and more problematic, conceptual problems with the model render it inapplicable in a large number of situations, and particularly for the definition of new products.

Of concern here is the application of ideal points to scales of preference rather than the probability of choice. Preference measures have a stability lacking in choice probabilities. For example, the addition of Pepsi Cola to a product set would not be expected to change significantly the preference score for Coke. By contrast, the probability of choosing Coke might change radically. There are two reasons for this change. First, the existence of any new product in the product set is likely to reduce the probability of purchase of all brands; and secondly, because of its similarity to Coke, the addition of Pepsi is likely to have greater effect on the choice of Coke than other less similar brands. Thus, choice probabilities depend not only on the distance of a stimulus from an ideal, but also on the positions of the other stimuli in the choice set. A preference model is a necessary stage in a choice model, but not sufficient in itself.

This paper considers several routines for determining product spaces and positioning individual ideal points within those spaces. Implementation problems are then discussed, followed by a consideration of the conceptual problems with the models. An attempt is made to show why ideal point models have not, and probably will not, be very effective modeling preference.

### Methods of Implementation

This section is limited to the feasibility of placing

individual ideal points in a space that applies across individuals. There are many ways to derive such a space. Following modifications of Coomb's (1964) unfolding paradigm, a joint space of stimuli and subjects can be derived from preference orderings. Alternatively, ideal points can be fitted into a space that already exists. Such spaces can have either a psychological or physical basis. Psychological spaces can be derived from similarity data using techniques such as TORSCA (Young and Torgerson, 1967), or INDSCAL (Carroll and Chang, 1970). Similar geometric representations can be produced through either the use of discriminant (Johnson, 1971) or principle components analysis (Morgan & Purnell, 1969) of the subjects' ratings of the stimuli by attributes. McKeon (1962) has shown that these apparently different methods are, under reasonable assumptions, theoretically equivalent. Furthermore, empirical work (Jones and Pessemer, 1974; Huber, 1973; Moore, 1974) indicates high correspondence between the spaces produced by alternate methodologies.

Several techniques are available to position ideal points in a defined space such that the closer a stimulus is to an ideal point, the more it is preferred. PREFMAP (Carroll, 1972) fits a quadratic function to the preference surface, the mode of which is the ideal point. The program assumes that preference is inversely proportional to the square of the distance from the ideal point. There are four phases to the program. Phase I allows the space to be rotated and differentially stretched for each subject. Phase II allows a differential stretching of the axes but no rotation. Phase III assumes no differential stretching and results in circular isopreference curves about the ideal points. Finally, Phase IV assumes a vector model which can be interpreted as placing the ideal point an infinite distance from the space. (See Green and Carmone, 1970, for a fuller explanation of this perspective.) The four phases are, in fact, progressively restricted quadratic regression, until in its final phase the coefficients of the squared terms are assumed to be zero and the result is simply linear regression. The phases form a linear hierarchy of models and the routine provides F-tests to ascertain whether the additional parameters of the more complex phases produce greater improvement in fit than could be expected by chance.

PREFMAP has a nonmetric option which allows the preference data to undergo a monotone transformation such that the fit of the linear version will be maximized. Under this option, the statistical tests comparing the fit of the models become descriptive at best. Further, unless one's preference data is metrically very weak, there is evidence (Huber, 1975) that the use of the nonmetric option reduces predictive accuracy simply because less of the available information is used.

A second routine that can be used to position ideal points in a space is LINMAP (Shocker and Srinivasan, 1974). It derives from the same assumptions as Phase II of PREFMAP: preference is inversely proportional to the square of the distance from the ideal point in a space that allows differential stretching of the axes for each subject. Instead of using quadratic regression to determine the position of the ideal points, LINMAP uses linear programming. This approach allows constraints to be put on the weights of the axes so that they are not negative. A negative weight implies the existence of an anti-ideal so that the modal point

of the quadratic surface is the least, rather than the most, preferred. The ability to avoid anti-ideals is advantageous in two senses: (1) ideal points with negative weights cannot be clustered with those with positive weights thus making segmentation by clustering ideal points more difficult, and (2) anti-ideal points result in the counter-intuitive result of increasing marginal returns--the more one moves away from an anti-ideal along a given dimension, the more important that dimension becomes relative to other dimensions.

A second potential advantage of LINMAP is that it estimates the position of an ideal point directly. By contrast, PREFMAP uses least squares quadratic regression to estimate the preference surface and then takes a ratio of estimated parameters to arrive at the location of the ideal point. Since the estimated parameters are not independent, their ratio is not unbiased. Furthermore, the position of the ideal point may be very sensitive to small changes in the denominator. This, in fact, occurs when the ideal point is positioned outside of the range of the stimulus space and the preference structure approaches that of a vector model. In such cases, the position of the ideal point has been found to be quite sensitive to small changes in the input data. More work needs to be done dealing with this problem and in assessing the relative effectiveness of the routines. Empirically, a comparison of the predictive effectiveness of LINMAP against PREFMAP would be valuable; while analytically, the statistical properties of the solutions of either of these routines have still to be worked out.

#### Problems of Implementation

The foregoing has discussed a number of ways to form a stimulus space and to position ideal points within the resultant space. This section discusses limitations in these algorithms and suggests procedures that will minimize these problems. In particular, it considers the assumption of symmetry of preference about the ideal point and the problem of multiple ideal points. These are what might be called "technical problems" in that it is possible to specify the form of a correct solution within the context of the ideal point framework. Later, conceptual problems with the ideal point model are considered. These are far less tractable.

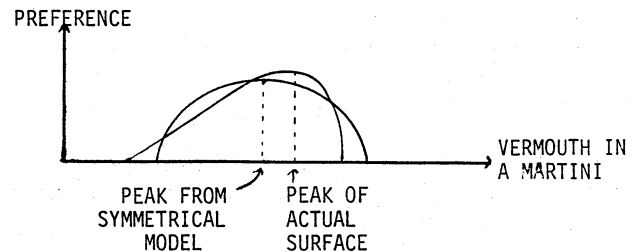
Virtually all schemes for positioning ideal points in attribute space assume that, with respect to any attribute, deviation from either side of the ideal point has equal affective consequences. Pessemer's (1972) model is a notable exception but has the difficulty of requiring from respondents explicit judgments of ideal levels on each attribute. The reason implicit positioning of ideal points has assumed symmetry stems from considerations of degrees of freedom and programming ease. Undirectional distance in a space of any dimensionality can be simply stored as one number. Directionality, on the other hand, requires attribute-by-attribute information since positive and negative directions only make sense within the context of an individual attribute. In terms of degrees of freedom, any model that allows differential stretching on either side of an ideal point requires at least one degree of freedom for each dimension.

One of the main problems with the assumption of symmetry is that it is so often counter-intuitive. Consider the ideal temperature for hot tea: being twenty degrees under the ideal point is certainly not equivalent to being twenty degrees over. One may be unpleasant, but the other is painful. Similarly, displeasure with vermouth in a martini might be far more sensitive to too much vermouth than too little. If symmetry is assumed, and it does not occur, systematic biases can occur in the location of the ideal point so that it no longer represents the most preferred point in the preference

distribution. The problem is analogous to the distinction between the mode and the mean in skewed probability distributions. In the hypothetical case given in Figure 1, the assumption of symmetry results in an ideal point with less vermouth than is most preferred. This divergence is most easily observed by the lack of fit if the symmetrical model is moved so that its mode coincides with the mode of the actual surface.

FIGURE 1

#### EFFECT OF SKEWED PREFERENCES ON SYMMETRICAL IDEAL POINT MODEL



There are several possible solutions to this problem. If few dimensions are used, the best way to determine an ideal point is simply to look for it. SYMAP (Shepard, 1970) is a computer routine that provides such a perspective in two-dimensional space. Designed to produce geophysical contour maps, the routine estimates the height (preference) as a function of the level and trend of the surrounding points. The effect is similar to two-dimensional smoothing adjusted for trend. The mode of the surface is easily determined by inspection.

In more than two dimensions, other techniques are required. One possibility is to modify a technique such as PREFMAP to weigh inputs to the quadratic regression in terms of their preference. This would mean that points closest to the ideal point would have the greatest weight in determining its position. The generated ideal point would then be closer to the mode of the response surface.

A second possibility involves first smoothing the preference data to reduce the effect of random fluctuations. The techniques of SYMAP generalized to many dimensions could be used for such a task. Once the random noise has been minimized, the method of steepest ascent or ridge analysis (see Myers, 1971) could be used to find the modal preference values for an individual.

Such a procedure might also be effective in determining whether there are multiple peaks within the space. Consider the choice of color for an automobile. Relatively small differences in hue, saturation and brightness can be expected to evoke different connotations and therefore, different preferences. Fairly specific shades of different colors might form local preference peaks. For a manufacturer offering a limited number of colors, it is important to estimate the distribution of these peaks, both within and across individuals. Local search routines on a smoothed preference surface could determine such local minima for individuals and these could then be aggregated across individuals.

To summarize, this section has dealt with the possibilities of asymmetry of preference and multiple peaks within a stimulus space. The notion of symmetry is critical to the ideal point model if one conceptualizes the ideal point to be the mode or peak of the preference distribution. Under such circumstances, routines can be modified to provide better estimates of the mode of the distribution. In the same way, the existence of multiple preference peaks, within a stimulus space, was shown to result in biased estimates unless an attempt is

made to account for the different peaks. The next section considers those conditions under which, even given the best possible implementation procedures, ideal point models cannot be expected to work.

### Conceptual Problems

The first requirement for an ideal point model--or any model-- to represent preferences in space, is that preferences be causally related to the underlying space. It is easy to think of examples where this would not be the case. Tucker (1964) shows that strong brand preference can develop for essentially identical brands of bread. There, the physical space would not predict preferences, while any psychological space would predict preferences only to the extent that a halo effect distorts attribute ratings (see Beckwith and Lehmann, 1975). For example, a psychological space might not be expected to predict the preferences of a man who buys a brand of automobile simply because his family has always bought that brand. Habit and familiarity are functions of the subject's past experiences rather than functions of the stimuli themselves. Where such components largely determine preference, a joint space of stimuli and subjects should not be used to specify preference for new brands simply because causality runs from usage to judged attributes rather than the other way around. Thus, any attribute space can be expected to be ineffective predicting preferences to the extent that preferences are based on individual habits or differential learning. Since most frequently purchased consumer goods fit into this category, it is perhaps not surprising that attributes have done such a poor job predicting preferences, except as a result of the halo effect.

Even given that the space is relevant, ideal point models have further problems. This section attempts to show that ideal point models are basically incompatible with vector models. Here, a vector model is one where preference never reaches a peak with respect to a given attribute. For example, more comfort is always preferred to less comfort, other things being held equal. An ideal point model, by contrast, assumes that beyond a certain point, preference falls with more of the attribute.

It is well-known that an ideal point model is equivalent to a vector model if the ideal point is an infinite distance away from the product set (Green and Carmone, 1970). Although the distance measure becomes undefined in such situations, the directional cosines of the lines from the centroid of the stimuli to the ideal do not. What is less well-known is that in a mixed mode situation, where some of the dimensions are monotone while others are singled peaked with respect to preference, both the distances and the directional cosines are undefined. This is not a problem empirically. In the quadratic regression of a routine such as PREFMAP, the isopreference curves are simply parabolas, and the probability of the denominator going to zero with real data is very low; but it does mean that the appealing geometric metaphor of the ideal point breaks down in this case. Furthermore, this consideration limits ideal points in the analysis of a very critical marketing variable, price. To the extent that a lower price is always preferred, price is monotone with preference; thus, the ideal point is undefined where price is one of the attributes.

If the foregoing argument is accepted, it implies that ideal point models should not be used where price, or any other vector attribute, is determinant. Ideal points might be appropriate in formulating the mix of ingredients in a candy bar or the ideal image for a cigarette since for such items price would not be ex-

pected to vary as a function of a brand's position in the product space. In such models, price is implicitly assumed to remain constant. Unfortunately, such products (e.g. candy bars and cigarettes) are likely to be those where learning and habit are important determinants of preference. To the extent that this problem occurs, neither the ideal point, nor any other spatial representation, is likely to be very useful.

### The Purchase Ideal: Ideal Point Within a Vector Framework

Having spoken of the incompatibility of joint vector and ideal point models, consideration must be given to a model of Pessemier's (1972) that derives ideal points in a space defined primarily by attributes which are monotone with respect to preference. This interior ideal point is defined by asking subjects to specify the level of attributes of the product they would most likely purchase. Such a purchase ideal is often an internal point in the attribute space reflecting the trade-offs and constraints that exist in actual purchase decisions. While such a model will result in approximate solutions, it will be shown that the divergence of the model from reality severely limits its use on many product sets and, in particular, for the evaluation of new products.

The purchase ideal is analyzed by positing a simple economic model of the purchase decision. It will then be possible to highlight those situations where the purchase ideal results in improper predictions. To make the discussion concrete, we shall consider choice among automobiles and assume that the dimensions of sportiness, luxury and price summarize all that is relevant in the purchase decision. Preference for purchase can be represented as:

$$\text{Preference} = F_1(\text{Sportiness, Luxury, Price}) \quad (1)$$

Here, it is assumed that the partial derivatives of  $F_1$  are always positive, though decreasing, for sportiness and luxury, and negative for price. Two models of the purchase ideal will be explored which account for Equation 1. The first of these positions the ideal in a space of only sportiness and luxury, while the second also includes price as an independent dimension.

FIGURE 2

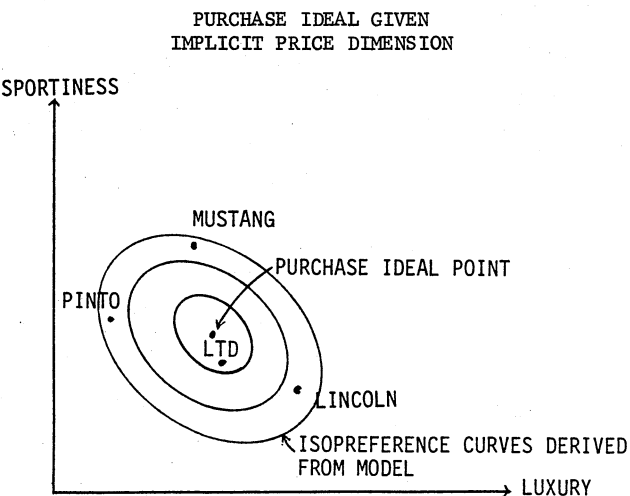


Figure 2 provides an illustrative purchase ideal for automobiles in a two-dimensional space of sportiness

and luxury. The internal ideal occurs at the point where the diminishing returns from improved sportiness and luxury are simply not worth the additional cost. Since price is not explicitly part of the Figure 2, this interpretation requires that price makes up an implicit third dimension which diminishes the overall value of a purchase as one moves beyond the purchase ideal. This dimension is contained in the dimensions of luxury and sportiness so that a given level of these attributes implies, on average, a given price. Or,

$$\text{Price} = F_2(\text{Sportiness}, \text{Luxury}). \quad (2)$$

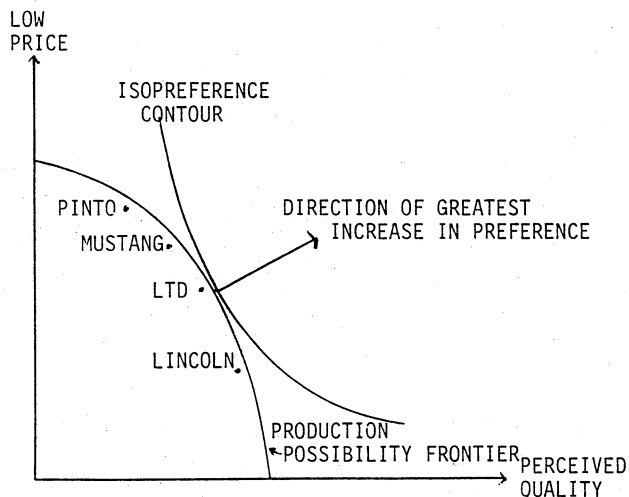
Equation 2 enables one to predict purchase preference in Equation 1 without using price. Its difficulty lies in the fact that even small error in Equation 2 can be expected to result in poor predictions. Consider a group of automobiles with approximately equal sportiness and luxury. These will have different prices which will strongly affect their probability of purchase. Yet, by the model in two dimensions, they would all be equally preferred for purchase. The implications for management might be even more misleading. Since the model assumes that a given level of attributes imply a certain price, any new product that does not conform to that implicit relationship will result in incorrect prediction. Specifically, a product with new lower price will not be conceived as superior to the original product with the same attributes.

This problem of causal misdetermination is general to many multivariate models. It occurs whenever one is working with a reduced space where (highly correlated) attributes are grouped in one factor. The predictions for a given product are only accurate insofar as that product is typical of the interrelations in the product set.

The problem can be mitigated, but not solved, by allowing each dimension to be independent of the others. In the automobile example, this can be illustrated by allowing price to be an independent dimension. For expositional purposes, sportiness and luxury are aggregated into a new dimension, "quality", and the decision is pictured as in Figure 3.

FIGURE 3

PURCHASE IDEAL GIVEN  
EXPLICIT PRICE DIMENSION



The production possibility frontier expresses the greatest perceived quality for a given price. The ideal

point is where this curve is just tangent to the indifference curve between quality and price. This point, while being the most preferred from the available set, is not the "ideal point" in the normal sense of the phrase. Points beyond the ideal are not less preferred, but merely perceived to be infeasible. Furthermore, given the economic model postulated, a routine such as PREFMAP would not determine an ideal at the point shown, but rather a vector indicating the direction preference increases most rapidly. While this vector is clearly the appropriate model of consumer behavior in this hypothetical situation, it should be pointed out that the high correlation between price and quality will lead to a degree of indeterminacy with respect to the preference vector. This can be avoided by providing subjects with hypothetical choices on both sides of the production possibility frontier so that the dimensions of price and quality are more nearly orthogonal.

To summarize, then, internal ideal points can result in a reduced space of vector attributes where correlated attributes with opposite affective consequences (such as price with sportiness and luxury) are combined. Such a purchase ideal can result in misleading managerial recommendations to the extent that the reduced space slightly misrepresents potential offerings. By making each dimension independent, the internal ideal is seen to be an artifact of the reduced space. In the full space, the vector model, constrained by the production possibility frontier, is the appropriate model which does not result in incorrect preference predictions for potential new products that are atypical of the present set.

#### Summary and Conclusions

The paper has dealt with four problems encountered with the representation of individual preference as an ideal point in a geometric stimulus space.

1. Lack of symmetry of preference about the ideal.
2. Multiple ideal points for an individual in a given space.
3. Preference due to subject-specific attributes such as habit or past behavior rather than product-specific attributes such as density, or masculinity of image.
4. Situations where some or all of the attributes are monotone with respect to preference.

The first two can be considered technical problems, at least it is possible to specify the kind of solution that is necessary to solve the problem. The second two are conceptual problems, and more serious. They place a restriction of the applicability of ideal point models. The first restriction involves the inapplicability of ideal point or any spatial model for those product sets where habit or past usage are important determinants of preference. Because of the halo effect, such models may fit existing products fairly well; but since causality runs in the wrong direction, predictions on new products will generally be wrong. The second restriction involves the incompatibility of the ideal point model with one or more vector attributes. This means that many important attributes, such as price or perceived quality, simply cannot fit into the ideal point paradigm.

With such restrictions, it is difficult to envision many situations where an ideal point model would be useful. One wonders whether, in the light of the last two problems, it is worth fixing the first two.

## References

N.E. Beckwith and Donald R. Lehmann, "The Importance of Halo Effects in Multi-Attribute Attitude Models," Journal of Marketing Research, 12(August, 1975), 265-275.

M. Braun and V. Srinivasan, "Amount of Information as a Determinant of Consumer Behavior Towards New Products," Working Paper 7303, (Rochester, New York, Graduate School of Management, University of Rochester, 1973).

J.D. Carroll and J.J. Chang, "Analysis of Individual Differences in Multidimensional Scaling Via an N-Way Generalization of Eckart-Young Decomposition," Psychometrika, 35(1970), 283-319.

J.D. Carroll, "Individual differences in multidimensional scaling," Multidimensional Scaling: Theory and Application in the Behavioral Sciences, (in R.N. Shepard, A.K. Romney, S. Nerlove, Editors, Academic Press, New York, 1972).

Clyde H. Coombs, A Theory of Data, (New York: John Wiley, 1964).

P.E. Green and A.K. Jain, "A Predictive Study in Perceptual and Evaluative Mapping," Unpublished paper, (University of Pennsylvania, 1972).

P.E. Green and F.J. Carmone, Multidimensional Scaling and Related Techniques in Marketing Analysis, (Boston: Allyn and Bacon, 1970).

Joel Huber, "The Psychophysics of Taste: Perceptions of Bitterness and Sweetness in Iced Tea," Association for Consumer Research 1973 Proceedings, (S. Ward and P. Wright, Editors, Urbana, Illinois, 1973), 166-181.

Joel Huber, "Predicting Preferences on Experimental Bundles of Attributes: A Comparison of Models," Journal of Marketing Research, 12(August, 1975), 290-297.

R.M. Johnson, "Market Segmentation: A Strategic Tool," Journal of Marketing Research, 8(February, 1971), 13-18.

W.H. Jones and E.A. Pessemier, "Single Subject Discriminant Configurations: An Examination of Reliability, Validity and Joint Space Implications," Institute Paper No. 451, (Krannert Graduate School of Industrial Administration, Purdue University, 1974).

James J. McKeon, "Canonical Analysis: Some Relations Between Factor Analysis, Discriminant Function Analysis, and Scaling Theory," Psychometric Monographs, 13(1962).

William L. Moore, "A Comparison of Product Spaces Generated by Multidimensional Scaling and by Single Subject Discriminant Analysis," Institute Paper No. 477, (Krannert Graduate School of Industrial Administration, Purdue University, 1974).

N. Morgan and J. Purnell, "Isolating Openings for New Products in Multidimensional Space," Journal of the Market Research Society, 2(July, 1969), 246-266.

Raymond H. Myers, Response Surface Methodology, (Boston: Allyn and Bacon, 1971).

E.A. Pessemier, "Stratop: A Model for Designing Effective Product and Communication Strategies," Institute Paper No. 470, (Krannert Graduate School of Industrial Administration, Purdue University, August, 1974).

E.A. Pessemier, "Multi-Attribute Models for Predicting Individual Preference and Choice," Institute Paper No. 346, (Krannert Graduate School of Industrial Adminis-

tration, Purdue University, 1972).

Donald S. Shepard, "SYMAP Interpolation Characteristics," Computer Mapping as an Aid in Air Pollution Studies, 2(Lab for Computer Graphics, Harvard University, Cambridge, Massachusetts, 1970).

Allan D. Shocker and V. Srinivasan, "A Consumer-Based Methodology for the Identification of New Product Ideas," Management Science, 20,6(February, 1974), 921-937.

W.T. Tucker, "The Development of Brand Loyalty," Journal of Marketing Research, 1(August, 1964), 32-35.

F.W. Young and W.S. Torgerson, "TORSCA, A Fortran IV Program for Shepard-Kruskal Multidimensional Scaling Analysis," Behavioral Science, 12(1967), 498.

AN ARGUMENT IN SUPPORT OF ORDINARY  
FACTOR ANALYSIS OF DICHOTOMOUS VARIABLES

Larry Percy, Ketchum, MacLeod & Grove

Abstract

The number of scale points necessary to effect a cogent factor analytic solution is explored. An argument is offered for the use of ordinary factor analysis in dichotomous variables by comparing the factor solutions generated from 2-point Likert scale and 5-point Likert scale representations of the same items. The results are discussed in terms of congruence and possible implication.

Introduction

One is frequently confronted in designing a research study with the disquieting dilemma of whether or not to use a multi-point scaling of a continuous variable and feel reasonably comfortable with the reliability and validity of the measurement, along with its applicability to most multivariate models, or to opt for a more convenient dichotomous variable and wonder if in fact one is somehow "losing" information, and quite possibly violating the assumptions of many multivariate models. Evidence is accumulating that would seem to arrest any continuing anxiety over this question.

On the problems of reliability or validity of measurement, neither are dependent on the number of scale intervals. Matell and Jacoby (1971), in a study based on a series of eighteen different scale alternatives varying from 2 through 19 scale intervals, report no significant difference evidenced in either reliability or validity, regardless of the number of scale intervals used. Similar results were obtained when higher ordered scales were collapsed. They conclude that the typical Likert-type scale measures primarily direction, and only marginally intensity. The overwhelming majority of the variance contained in a measurement derived from Likert-type composite rating scales is accounted for by the directional component; so increasing the number of scale intervals does not increase the precision of the instrument. One is not "losing" information, accordingly, by employing a dichotomous scaling

The second question is more widely debated: what is the effect on a multivariate model of reducing a continuous measurement scale to a dichotomy, particularly when the Pearson product-moment correlation is used to represent the relationship between the variables? Addressing a specific interest in factor analysis and the so-called problem of the appropriate correlation measure, one may find support for any of his favorite coefficients: Carroll (1961) feels the tetrachoric is perhaps more suited when dealing with dichotomous variables, while Horst (1965) states that its use is unwarranted; Cattell (1952) favors phi-over-phi-max. Henrysson and Thunbert (1965) discuss the tetrachoric vs. phi coefficients in factor analysis. In one of the few empirical studies (as reported in Rummel, 1970) addressing itself to this question, Comrey and Levonian (1958) concluded that the "same" factors result from all three, and thus recommended phi (or the usual Pearson product moment correlation associated with most factor analysis programs) as the most reasonable choice. Their reasoning followed because

Phi-over-phi-max and the tetrachoric lead to overly high communality estimates.

After coming 'round the bush, so to speak, in determining the appropriateness of actually using the Pearson product moment correlation with dichotomous data, what actual differences in factor solutions are to be expected as the number of scale intervals of a continuous variable is reduced to two? Martin et al. (1974) hypothesized that the reduction in the number of scale intervals would have a substantial effect on factor analysis results. They were wrong. Utilizing seven different numbers of scale intervals down to a dichotomy, a general decrease in size of the eigenvalues, communalities, and factor loadings was noted as the number of scale divisions was reduced; however, the patterns of the rotated factor loadings were not appreciably affected, and remained consistent over the entire range of scale intervals. So much for the effect of the number of scale intervals on the results of a factor analysis.

Yet in the typical application of factor analysis in marketing research, the magic number of scale intervals remains five, along with severe warnings to those foolish enough to suggest only two. Guertin and Bailey (1970), for example, have stated that it is highly undesirable to use less than five interval scale points, and that to go beyond five is not particularly useful in providing more information. It is the point of this paper to debunk such notions.

Specifically, this paper investigates the effect on a factor solution of collapsing a typical 5-point agreement scale into a dichotomous measure. From a pragmatic standpoint we wish to see if the dichotomous scale provides sufficient information to uncover the latent dimensionality of a series of attitude questions as efficiently as a 5-point scale. If such is the case, from a marketing research viewpoint the saving in questionnaire administration time and the increase in respondent motivation to complete a long list of scales occasioned by the reduction in the number of scale intervals, would suggest that dichotomous scales substitute for the traditional 5-point scale, particularly in the development phase of an inquiry where intensity of response is conceptually less compelling than direction.

Study Design

The design of this study involves the utilization of data already in hand measuring homemakers' attitudes toward the meal along 20 variables. These variables resulted from a prior factor analysis of a much greater number of variables. Each of the 20 attitude variables are presented to subjects as a randomly shuffled deck of cards, each card containing a single variable. The subjects are then asked to sort the cards into appropriate piles on a sorting board reflecting a typical 5-point Likert-type agreement scale: five blocks labeled Agree Strongly, Agree Slightly, Neither Agree nor Disagree, Disagree Slightly, and Disagree Strongly. These data represented a national probability sampling of homemakers.

Plan of Analysis

The data collected are treated as a traditional 5-point Likert scaling, and are then collapsed as follows to dichotomous variables:

1. Subjects indicating they Agree Strongly or Agree Slightly with a variable are rescored as Agree for that item.
2. Subjects indicating they Disagree Slightly or Disagree Strongly with a variable are rescored as Disagree for that item.
3. Subjects indicating they Neither Agree nor Disagree with a variable are independently introduced to a program which randomly assigns them as Agree or Disagree for that item.

Both sets of data are then submitted to an R-mode factor analysis and the resulting solutions compared for congruence. This provides a comparative analysis of a 5-point Likert-type scale measurement with a dichotomous scale measurement over a common set of variables. The assumption is that both scales measure the same information, are submitted to the same factor analytic procedure, and should thus yield comparable factor solutions.

Results of the Analysis

The factor solution for the variables as measured by the 5-point Likert scale revealed eight latent dimensions of attitude (restricting rotation to those factors associated with eigenvalues greater than unity).

The resulting rotated factor loadings are detailed in Table 1. An interpretable solution of the eight dimensions according to appropriate criteria is shown in Table 2.

TABLE 2

Factor Interpretation of the Eight Dimensions Measured by the 4-Point Likert Scale

FACTOR	VARIABLE	LOADING
One	2	-.70770
	20	-.68653
	3	-.68488
Two	4	.72000
	10	.48354
Three	13	.69655
	14	.60378
Four	11	-.70533
Five	1	.74757
	7	.67372
Six	19	-.81462
	16	.58526
Seven	15	-.84709
Eight	9	-.77827

TABLE 1

Rotated Factor Matrix of the Eight Dimensions Measured By the 5-Point Likert Scale

VARIABLE	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	FACTOR 7	FACTOR 8	h <sup>2</sup>
1	-0.05793	-0.06037	-0.05056	0.07799	0.74757	0.16102	0.01899	0.00705	.60
2	-0.70770	0.10200	-0.04550	0.00439	0.01159	0.01152	-0.16774	-0.11729	.56
3	-0.68488	0.00818	0.05534	-0.06533	0.06891	0.04867	-0.02977	0.27125	.56
4	-0.01945	0.72000	0.11119	0.01008	-0.13215	0.05727	0.00766	0.06529	.56
5	-0.30960	0.15273	0.38171	-0.00130	0.14476	-0.14537	0.12756	0.04247	.32
6	-0.20640	0.32630	0.08121	0.28178	0.40342	-0.07560	0.09604	-0.06635	.42
7	-0.06782	-0.04858	0.10826	-0.17633	0.67372	-0.10288	-0.07232	0.11292	.53
8	-0.21661	0.21944	-0.08663	0.02672	-0.02639	0.12221	0.11896	0.46439	.35
9	-0.04073	-0.03243	0.06126	-0.05611	0.14544	-0.07306	-0.07470	0.77827	.65
10	0.32211	0.48354	-0.03041	-0.13545	0.12614	-0.08770	-0.18705	0.21069	.46
11	-0.00513	-0.03141	0.12779	-0.70533	0.02972	-0.05223	-0.01279	-0.05048	.52
12	-0.08133	0.32722	-0.32294	-0.58496	0.06827	0.07828	0.08334	0.08666	.58
13	-0.10124	-0.14338	0.69655	0.06610	-0.02705	0.21440	0.01568	0.08962	.58
14	0.12455	0.24759	0.60378	-0.13503	0.06346	-0.02289	-0.02171	-0.15390	.49
15	-0.05057	-0.00594	-0.00466	0.01528	-0.02060	0.01035	-0.04709	-0.01057	.72
16	-0.02514	-0.13730	0.26450	-0.15165	-0.11213	0.58526	-0.25131	0.16042	.56
17	-0.13259	0.26264	-0.13509	0.41375	0.18185	-0.05440	-0.41843	-0.03830	.49
18	-0.42101	-0.21418	0.08323	-0.17093	0.16078	0.13236	-0.05379	-0.40714	.47
19	-0.05523	-0.11136	0.03829	-0.09207	-0.11843	-0.81462	-0.14563	0.12526	.74
20	-0.68653	-0.13580	0.03115	0.08693	0.09820	-0.11038	0.06221	0.10884	.54
Eigenvalue	2.21805	1.55044	1.39508	1.25720	1.15652	1.05696	1.03957	1.01030	

Each variable contributing approximately one-quarter of the variance in any one factor (i.e. a rotated loading with an absolute value greater than .5) was considered in interpreting the meaning of that factor. Additionally, any marginal variable with a rotated loading of absolute value greater than .4 but less than .5 was considered if the sense of the variable was compatible with those loading more strongly. One will notice in checking the rotated loadings with those making their way into the solution that certain variables which meet the criteria outlined above nevertheless are not listed (e.g. variable 12 which loads -0.58496 on factor 4 but is not shown in the interpretation). These occasional stray variables find their way into the solution space because of a mathematical congruence with the factor rather than any sense of psychological homogeneity and must be dismissed by the analyst as noise in an otherwise condign system.

A second factor solution for the variables, this time as measured by the dichotomous scale, was sought in a manner identical to that of the initial factoring of the 5-point Likert scale items. This solution also revealed eight latent dimensions of attitude. The resulting rotated factor loadings for this second solution are detailed in Table 3, with the corresponding interpretable solution of the eight dimensions shown in Table 4. Again, the same procedure was used in determining which variables describe the latent attitude dimensions.

From a visual inspection of the two factor patterns, as decocted in Table 2 and Table 4, it appears quite obvious that a great deal of similarity exists between the two solutions. One may conclude that factor one, as derived from the first solution, is all but "identical" to factor one as derived in the second

TABLE 4

Factor Interpretation of the Eight Dimensions Measured by the 2-Point Likert Scale

FACTOR	VARIABLE	LOADING
One	3	.65288
	2	.63705
	20	.63537
Two	4	-.63807
	10	-.61561
Three	19	.75614
	16	-.50457
Four	1	.70706
	7	.62717
Five	15	.68776
Six	14	.61251
	13	.57135
Seven	11	-.82312
Eight	17	-.62312
	6	-.51883

TABLE 3

Rotated Factor Matrix of the Eight Dimensions Measured by the 2-Point Likert Scale

VARIABLE	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4	FACTOR 5	FACTOR 6	FACTOR 7	FACTOR 8	h <sup>2</sup>
1	0.09124	0.05299	-0.24558	0.70706	-0.04330	-0.10723	0.10145	-0.10060	.60
2	0.63705	0.12777	0.01867	-0.08485	0.00692	-0.02078	-0.04691	-0.14192	.45
3	0.65288	-0.03977	-0.03096	0.12306	0.03417	0.05082	-0.06987	0.08878	.46
4	0.00116	-0.63807	-0.15303	-0.10705	0.02112	0.12287	-0.16307	-0.12129	.50
5	0.28279	-0.00839	0.17624	-0.03052	-0.16776	0.50447	0.02330	-0.16732	.42
6	0.09860	-0.02239	-0.00431	0.35124	-0.05884	0.01592	0.03653	-0.51883	.41
7	0.07380	-0.09329	0.17051	0.62717	0.04915	0.15297	-0.26010	-0.00945	.53
8	0.25237	-0.24959	-0.28475	-0.26872	-0.07430	-0.26898	-0.16269	-0.17168	.41
9	0.29981	-0.37669	0.14068	0.21245	-0.02007	-0.00811	0.33091	0.39281	.56
10	-0.13512	-0.61561	0.01531	0.13771	-0.02309	-0.01597	-0.03843	-0.08285	.42
11	0.00295	-0.07160	0.09648	0.07974	-0.07496	-0.07383	-0.82312	0.12883	.73
12	0.16784	-0.14275	0.03564	0.00598	-0.67685	-0.12943	-0.01359	0.17325	.55
13	0.15824	0.00159	-0.15621	-0.04396	0.30531	0.57135	-0.19486	0.15245	.53
14	-0.12596	-0.04088	-0.08272	0.06796	-0.10056	0.61251	0.12635	0.03336	.43
15	0.13990	-0.09527	0.08502	0.00448	0.68776	-0.22142	0.06761	0.07742	.57
16	0.13425	0.00064	-0.50457	-0.03314	0.07671	0.06993	0.03366	0.40099	.45
17	0.12180	-0.18867	0.04367	-0.04323	0.14736	-0.02922	0.12820	-0.62312	.48
18	0.20440	0.49771	-0.23641	0.07794	0.05184	0.08145	-0.25276	-0.06047	.43
19	0.08755	-0.01091	0.75614	-0.07472	0.06498	-0.03650	-0.06439	0.04818	.60
20	0.63537	0.14449	-0.00589	0.11669	-0.03758	0.04191	0.11723	-0.05969	.46
Eigenvalue	1.81779	1.43237	1.24929	1.18831	1.13413	1.10965	1.05885	1.01264	



solution. Similar conclusions may be drawn for all but one factor in each solution; factor two matches factor two; factor three matches factor six; factor four matches factor seven; factor five matches factor four; factor six matches factor three; factor seven matches factor five. Only factor eight in each solution is unique.

In the first solution, factor eight represents a single variable, which barely increases the accounted for variance; the same is true of the two variable factor eight in the second solution. It is probable that a seven factor-solution was in fact the most parsimonious. Otherwise, the solutions appear congruent regardless of whether the original input resulted from a 5-point or dichotomous scale (see Table 5).

TABLE 5

Comparison of the Two Factor Interpretations

5-Point Likert Scale Factor Solution		2-Point Likert Scale Factor Solution	
FACTOR	VARIABLE	FACTOR	VARIABLE
One	2	One	3
	20		2
	3		20
Two	4	Two	4
	10		10
Three	13	Six	14
	14		13
Four	11	Seven	11
Five	1	Four	1
	7		7
Six	19	Three	19
	16		16
Seven	15	Five	15

Although by inspection one might be satisfied that there was little real difference in the results of the two separate analyses, it would certainly help if there were some analytic way of comparing the goodness of fit between the two solutions. Rummel (1970) has pointed out that in addition to an intuitive mode of comparison two systematic models exist for comparing factors from separate analyses: a matrix comparison technique that transforms the factor solution of one study to a least squares fit to the factors of another; and a vector comparison method such as correlation or congruence coefficients.

Harman (1967) and Cattell (1966) both discuss a coefficient of congruence for determining the extent of similarity or agreement between factors obtained in different solutions. This coefficient, as developed by either Tucker (1951) or Burt (1948) is precisely the same, and is given in matrix notation by Cattell as:

$$r_c = \frac{\sum XY}{\sqrt{\sum X^2 \sum Y^2}} \quad (1)$$

where X and Y are loadings of the same variable, not deviations of loadings. Harman's notation shuns matrix form and is given by:

$$\phi_{pq} = \frac{\sum_{j=1}^n 1^a_{jp} \cdot 2^a_{jp}}{\sqrt{\sum_{j=1}^n 1^a_{jp} \sum_{j=1}^n 2^a_{jp}}} \quad (2)$$

While this coefficient may appear to be similar in form to the product-moment correlation coefficient, notice that (a) one is not dealing with deviations from loadings, but with actual loadings, and (b) one is summing over common variables not individuals. Any attempt to merely correlate the columns of the factor matrices would be insufficient for one would be ignoring the difference between the origins and the differences of mean loading. Both procedures suffer from a lack of attention to possible differences in level between two patterns, but for purposes of this analysis level is not a crucial consideration. Were level an important consideration, the Cattell et al. (1966) pattern similarity coefficient  $r_p$  would avoid this difficulty by indicating a high congruence of loadings only when both the shapes of the patterns involved and the level are similar.

Because the n variables are identical for the solutions at hand, a straightforward application of equation (1) or (2) to the corresponding loadings of the hypothesized matching factors as shown in Table 5 is possible. Considering the factors derived from the 5-point Likert scale measurement as the first group X and the factors derived from the dichotomous scale measurement as the second group Y, coefficients of congruence were computed for each of the seven factor pairs. The resulting coefficients are shown in Table 6.

The coefficient of congruence can range in value from +1 for perfect agreement through zero for no agreement at all to -1 for perfect inverse agreement. It differs from a correlation coefficient, however, in not equating means (i.e. in not standardizing the data). The coefficients of congruence computed between these two solutions all indicate almost perfect agreement.

Harman (1967) points out that in situations where factors from two solutions may be matched visually, and the number of variables common to the factors is small, it can be expected that the coefficients will be high. It was assumed in these calculations that only those variables common to the factor interpretations were common to any two factors being compared. Table 6 illustrates the variables and their corresponding loadings which were used in the computation.

This whole notion of factor matching for congruence, is of course, consistent with the model of factors as determiners and with the more subtle idea that a factor either operates fully upon a variable or leaves it entirely alone. Any fall from a unit loading by a variable within a factor is purely a function of the

other factors that happen to be operating. When one thinks of a lower factor loading implying a smaller effect on a variable he is tacitly recognizing that for a particular situation other factors are more dominant.

TABLE 6  
Coefficients of Congruence

Factor Pair	Rotated Loadings <sup>a</sup>		Coefficient
	5-Point Scale Solution	2-Point Scale Solution	
1, 1	0.70770	0.63705	0.99976
	0.68653	0.63537	
	0.68488	0.65288	
2, 2	0.72000	0.63807	0.98455
	0.48354	0.61561	
3, 6	0.69655	0.57135	0.99436
	0.60378	0.61251	
4, 7 <sup>b</sup>	0.70533	0.82312	1.00000
5, 4	0.74757	0.70706	0.99996
	0.67372	0.62717	
6, 3	0.81462	0.75614	0.99940
	0.58526	0.50457	
7, 5 <sup>b</sup>	0.84709	0.68776	1.00000

<sup>a</sup>Absolute values

<sup>b</sup>Coefficient values for single variable factors which match are of course 1.00000.

If a factor exists and is central to the domain of study, it will occur regardless of the factor technique or instrumentality of measurement. The degree of efficacy in a factor concept is only reinforced by its appearance under varying sampling situations and forms of analysis. In this case, removal of intensity within the attitude measurement device did not alter the factor solution.

#### Discussions and Conclusions

Matell and Jacoby (1971) have indicated that reliability and validity are found to be independent of the number of scale points contained in a Likert-type rating scale. They additionally demonstrate that regardless of the number of steps originally used in collecting the data, re-scoring of these multi-point response scales to dichotomous measure does not result in any significant decrement in reliability or validity. They contend, as both Peabody (1962) and Cronbach (1950) have, that Likert-type scales measure primarily direction and only to a minor degree intensity.

Since any Likert scale measurement is concerned primarily with direction, utilization of a 2-point Likert scale realizes largely the same information as a multi-point Likert scale. Correlations of these data will reflect this fact because the basic monotone

relationship has not been altered. By observing normal cautions with the data, there is no meaningful effect on the correlation matrix transformation as a result of the number of scale points utilized.

Proceeding from these conclusions, a study of the congruence of the factor solutions of data as measured in a large scale study by both a 5-point Likert scale and a rescored 2-point Likert scale was undertaken. It was hypothesized that if neither the reliability or validity of one's data was affected by the number of points in a given scale, then the identity of the factor structure of influences that may be unearthed by the factor analysis would be manifest in the data regardless of the manner of collection. Although Nunnally (1967) argues that it is a mistake to compare factors in terms of the patterns of loadings in two analyses because the loadings are not the factors, a coefficient of congruence was none-the-less chosen as the measure of comparison due to its overwhelming prominence in the literature (cf. Cattell, 1966; Harman, 1967; Horst, 1965; Martin et al., 1974; Rummel, 1970). The evidence of the coefficients of congruence in the present study indicates that the ability to extract the underlying factor definition of a space proceeds independently of the number of scale points used for the Likert-type items.

The effect of these findings mean greater flexibility in study design. Data which were collected previously with different rating formats, or data within a given instrument collected with a variety of Likert-type items, may be collapsed into dichotomous measures and factor analyzed because of their now common base measure. It would even be possible to dichotomize open-ended responses along some agreement criteria and include them along with more ordered ordinal or metric measures in the factor model.

These results suggest a much more wide-ranging approach to the traditional uses of factor analysis. The structure of scale imposition no longer seems as critical. At the very least the present findings indicate that for efficiencies in large-scale collection a 2-point Likert scale may be effectively utilized. The resulting factor analysis of this large scale data collection could then form the basis of the critical questioning in a follow-up phase where intensity would be measured along a multi-point scale, but only for the relevant factors.

#### References

- Cyril Burt, "The Factorial Study of Temperamental Traits," British Journal of Psychology (Statistical Section), 1 (1948), 178-203
- \_\_\_\_\_, "Factorial Analysis of Qualitative Data," British Journal of Psychology (Statistical Section), 3 (1950), 166-85.
- J. B. Carroll, "The Nature of the Data, or How to Choose a Correlation Coefficient," Psychometrika, 26 (1961), 347-72.
- Raymond B. Cattell, Factor Analysis: An Introduction and Manual for the Psychologist and Social Scientist. New York: Harper and Row, 1952.
- \_\_\_\_\_, "The Meaning and Strategic Use of Factor Analysis," in Raymond B. Cattell, ed., Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally & Company, 1966.

Raymond B. Cattell, Malcom A. Coulter and Bien Tsujoka, "The Taxonomic Recognition of Types and Functional Emergents," in Raymond B. Cattell, ed. Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally & Company, 1966.

A. L. Comrey and E. Levonian, "A Comparison of Three Point Coefficients in Factor Analysis of MMPI," Educational and Psychological Measurement, 18 (1958), 739-55.

L. J. Cronbach, "Further Evidence on Response Sets and Test Design," Educational and Psychological Measurement, 10 (1950), 3-31.

Wilson H. Guertin and John P. Bailey, Jr. Introduction to Modern Factor Analysis, Ann Arbor: Edwards Brothers, Inc. 1970.

Harry Harman, Modern Factor Analysis (second edition). Chicago: The University of Chicago Press, 1967.

S. Henrysson and P. Thurnberg, Tetrachoric or Phi Coefficients in Factor Analysis. Report No. 27, Department of Psychology, University of Uppsala, Sweden, April, 1965.

P. Horst, Factor Analysis of Data Matrices. New York: Holt, Rinehart and Winston, 1965.

Jacob Jacoby and Michael S. Matell, "Three-Point Likert Scales Are Good Enough," Journal of Marketing Research. 8 (November, 1971).

Warren S. Martin, Benjamin Fruchter and William J Mathis, "An Investigation of the Effect on the Number of Scale Intervals on Principal Components Factor Analysis," Educational and Psychological Measurement, 34 (1974), 537-45.

Michael S. Matell and Jacob Jacoby, "Is There an Optimal Number of Alternatives for Likert Scale Items? Study I: Reliability and Validity," Educational and Psychological Measurement, 31 (1971), 657-74.

J. C. Nunnally, Psychometric Theory. New York: McGraw Hill, 1967.

D. Peabody, "Two Components in Bipolar Scales: Direction and Extremeness," Psychological Review, 69 (1962), 65-73.

R. J. Rummel, Applied Factor Analysis, Evanston: Northwestern University Press, 1970.

Ledyard R. Tucker, A Method for Synthesis of Factor Analysis Studies, Personal Research Section Report, No. 984. Washington, D.C.: Department of the Army, 1951.

Charles Wrigley and Jack O. Newhaus, The Matching of Two Sets of Factors, Contract Memorandum Report, A-32. Urbana, Illinois: University of Illinois, 1955.

A MULTIVARIATE TEST  
OF CAD INSTRUMENT CONSTRUCT VALIDITY

Michael J. Ryan, The University of Alabama  
Richard C. Becherer, Wayne State University

Abstract

Cohen's CAD instrument was examined for two specific validation issues through the use of principal components and factor analysis. Although the instrument did not meet the validation requirements, the results indicate that a firm foundation has been established for future development that may lead to a valid instrument.

Introduction

A common criticism of consumer research is the use of unvalidated instruments to generate data (Kollat, Engel, and Blackwell, 1970; Robertson and Ward, 1972). Until recently (Davis, 1971; Becherer, Bibb, and Riordan, 1973; Heeler and Ray, 1974; Horton, 1974) research investigating the crucial problems of test reliability and validity have been absent from the marketing literature. In a review of attempts to relate purchasing behavior to personality, Kassirjian (1971) concluded that measures commonly employed were generally intended for other applications and inadequate for the intended purposes. He suggested that consumer behavior researchers develop and validate their own instruments to measure the personality variables that go into the purchase decision. In line with this thinking, Cohen (1967, 1968) developed the CAD scales, an instrument designed specifically to diagnose interpersonal orientations useful in predicting and explaining consumer behavior.

Some success has been achieved in the initial applications of the instrument (Cohen, 1967; Cohen, 1968; Kernan, 1971) and both Kassirjian (1971) and Heeler and Ray (1974) in their inclusive reviews of the literature have cited the CAD as worthy of further investigation. While admittedly picking and choosing from the data, Cohen's (1967, 1968) initial published research linked each of the basic response orientations to specific patterns of product or brand usage as well as television and magazine preferences. Other encouraging results with the CAD were evidenced by Kernan's (1971) findings which suggest relationships between compliant, aggressive, and detached orientations and such things as information source utilization and fashion approval.

These findings must be tempered, however, by the fact that more "positive" results have not been published using the CAD scales. Further, research must be initiated to determine if the successful applications of the instrument are merely artifacts of invalid and unreliable measures. In an attempt to construct something close to a multitrait-multimethod matrix, Heeler and Ray (1974) view their limited CAD validation effort as positive and suggest that the CAD should be examined more extensively. The research reported here continues the validation process by examining an aspect of the instrument's construct validity.

The CAD Instrument

Cohen postulated that Horney's (1945) tripartite paradigm would provide a general model of interpersonal response traits useful in predicting a broad range of

product purchase decisions. Horney theorized that individuals had three predominant orientations. These were described as compliant, aggressive, and detached.

The compliant individual wants to be part of the activities of others. He wants to be loved, wanted, appreciated, and needed. He sees in other people the solutions to many problems of life, and wants to be protected, helped, and guided. This person tends to be oversensitive to the needs of others, overgenerous, overgrateful, and overconsiderate. This individual tends to avoid conflict and subordinates himself to others. Important attributes are goodness, sympathy, love, unselfishness, and humility.

The aggressive individual wants to excel, to achieve success, prestige, and admiration. Other people are seen as competitors. This type of person strives to be a superior strategist, to control his emotions, and bring his fears under control. Strength, power, and unemotional realism are seen as necessary qualities. People are valued if useful to one's goals. This person seeks to manipulate others, and will go out of his way to be noticed, if such notice brings admiration.

The detached individual wants to put emotional "distance" between himself and others. Freedom from obligations, independence, and self-sufficiency are highly valued. This type of person does not want to be influenced, or share experiences. Intelligence and reasoning are valued instead of feelings and conformity is disliked. This individual considers himself unique, possessing certain gifts and abilities that should be recognized without going out of his way to show others. Generally, this individual is somewhat distrustful of others.

As a first step in applying this notion, Cohen developed a 35 item instrument composed of three sets of scales designed to measure Compliant, Aggressive, and Detached interpersonal orientations. Each of the 35 items was followed by a six place bipolar adjective set ranging from "extremely desirable" to "extremely undesirable." These items were scored from one to six, with ten items included for each of the Compliant and Detached scales ( $C_1$  through  $C_{10}$  and  $D_1$  through  $D_{10}$ ) and fifteen items included on the Aggressive scale ( $A_1$  through  $A_{15}$ ). Each scale is scored by summing across items, with high scores (those at least one standard deviation above the sample mean for a trait) indicating the respondent's orientation toward one of the three groups. The individual questions are reproduced in the Appendix.

The Problem of Construct Validity

Since the purpose of the CAD instrument is to measure three separate constructs, it should have demonstrated validity for this purpose. That is, it should have construct validity. Nunnally (1967) outlines three major steps in this process: 1) specifying the domain of observables; 2) determining to what extent all, or some, of these observables correlate with one another; and 3) determining whether or not one, some, or all measures of such variables act as though they measured the construct.

The first step generally involves defining the construct as it relates to words at a lower level of abstraction. Horney initiated this step and Cohen has employed some of these words to construct the test items. Cohen (1968) has also furnished evidence for this aspect in terms of face validity.

The third aspect is a matter of empirical testing to determine if operationally defined constructs are related to valid measures of other similar constructs and do not relate to valid measures of dissimilar constructs. Cohen (1967) has also furnished evidence that the instrument meets this condition through use of correlational analyses.

The second condition concerns the adequacy with which the domain of observables fit together. That is, do the instrument items relate to each other in a manner consistent with the proposed relationships between the constructs they represent. In an indirect manner, Cohen (1967, 1968) and Kernan (1971) may have furnished some support for this condition. That is, their successes indicate that the measures have operated like the constructs that they represent were expected to operate. This may indirectly indicate that the items fit together. However, in other research the measures have not operated as expected (Cohen and Golden, 1972). Consequently, a direct test is needed.

Two demonstrations are needed to furnish direct evidence that the instruments fulfill the second condition. First, the items composing each of the three scales should correlate highly with one another. This indicates that the items in a particular scale all measure much the same thing. Second, the items composing the total instrument should split into three groups such that members of each group correlate highly with one another and correlate much less with the members of other groups. This indicates that three distinct things are being measured. Strictly speaking, the issues so far addressed concern reliability rather than validity since the 'things' referred to have not yet been empirically identified. To the extent that empirical groupings are congruent with those conceptually specified on an *a priori* basis, the measurements take on meaning. The first vestige of construct validity is evidenced when measures shown to be reliable are congruent with theoretical expectations. In the present case, if items *a priori* specified to indicate a construct cluster together but not across groups, the items show evidence of indicating one of the three construct orientations. This study addresses these issues through the use of principal components and factor analysis.

#### Methodology

The CAD questionnaire was administered to a sample of 175 undergraduate junior and senior college students located at two midwestern universities. Since reliability is a necessary prerequisite to validity testing, this aspect of the instrument was examined before proceeding. Cronbach's (1951) coefficient Alpha measure of internal consistency was utilized as a reliability estimate. The reliability estimates were:

Compliant Scale	.724
Aggressive Scale	.680
Detached Scale	.514

These estimates were low enough to attenuate correlations (Nunnally, 1967: 226). This raised the question of whether the scales were unreliable to the extent of hampering the instrument's use in a correlational study.

The size of the item intercorrelation matrix made it difficult to analyze within and across scale correlations. Consequently, the general factor analysis model was employed to resolve the set of 35 items (*p*) linearly in terms of a small number of factors (*m*) that could ultimately be identified as indicating the three orientations in question. This model assumes that the total variance of an item is composed of three parts: common, specific, and error variance. Specific and error variance are not separated and are referred to as unique variance. Error variance is assumed to be "unreliable" or random variance which is sample specific for a given item. Common and specific variance are assumed to represent "reliable" or systematic variance likely to be stable from sample to sample. Common variance is shared among items whereas specific variance is item specific.

The possibility of instrument unreliability had implications for choosing principal components or factor analysis as the method of obtaining the linear reduction. Principal components analysis (a factor solution of a correlation matrix with 1's on the diagonal) is the method which analyzes the total variance of the items. In matrix form, the principal components model appears as follows:

$$R = F F' \\ p \times p \quad p \times m \quad m \times p$$

where *R* is the correlation matrix and *F* is a matrix of factor loadings. In this case the linear resolution contains all the variance in the correlation matrix.

Since the scales have demonstrated low reliabilities (and therefore low common variance), submitting them to principal components analysis may have led to two general problems. First, if the majority of the variance in the item correlation matrix had been random, it would have been unlikely that variables would have clustered in the factor matrix. Second, if clusters had formed, the factors may have represented correlated error variance. Such factors would have been unlikely to replicate in another sample.

Methods of factor analysis, employing estimates of common variance in the principal diagonal of the correlation matrix, analyze only common variance (that portion of the reliable variance of an item which correlates with other items in the matrix). In matrix form, the factor analysis model appears as follows:

$$R = F F' + U \\ p \times p \quad p \times m \quad m \times p \quad p \times p$$

where *R* and *F* are as previously described except that *R* contains communality estimates on the diagonal instead of 1's, *F* only accounts for common variance, and *U* is a diagonal matrix of unique variances. Since error variance (in this case instrument unreliability) is a subset of unique variance it is removed from the *F* matrix. That is, the correlational analysis in the linear solution does not consider unreliability.

Since common variance is a subset of reliable variance, the factor analysis method should have produced stable factors and allowed clusters to form irregardless of error magnitude. This method could have been misleading, however, since the factors and clusters it would have produced would not indicate whether the instrument may be ineffective due to unreliability (large measurement error). This would have been indicated by failure to find clusters in the principal components analysis. Consequently, both principal components and factor analysis were employed in this research.

Estimated communalities in the factor analysis were the squared multiple correlations of that item with all other items in the matrix. This method was chosen since it is the lower bound estimate of common variance (Guttman, 1956) and is the most conservative and widely used of the communality estimates.

Following the Kaiser (1960) criterion, only factors with factor contributions (eigenvalues) of 1.0 or greater were retained in the principal components and factor analysis. Although there is some dispute about the appropriateness of this criterion, it was adopted since it is the most widely known and used decision rule for the number of factors problem.

In order to locate the test items with respect to the factors and obtain a stable representation of this location, retained factors were rotated to simple structure (Thurstone, 1947). Since Horney's notion and Cohen's operationalization indicated that the resulting rotated factors would be defined by items related to each other and not to other variables in the matrix, an orthogonal rotation was employed. The Varimax (Kaiser, 1958) solution to the orthogonal rotation was employed since it has proven to yield a good approximation of simple structure.

### Results

The principle components factor matrix is presented in Table 1 and the factor analysis is presented in Table 2. Six factors were retained in the principal components solution only one of which accounted for more than 10% of the total instrument variance. The seven factors together only accounted for 39.8% of the total instrument variance. The rotation did not produce simple structure and there were no factor loading patterns evident among the items. That is, the items did not separate into clusters associated with specific factors.

The factor analysis solution produced three rather clearly defined and one extraneous factor. These four factors accounted for 70.3% of the common variance in the instrument. The first three factors accounted for 60.8% of the common variance with each of the first two factors accounting for more than 22% and the third accounting for more than 15% of the total common variance.

The factor loadings indicated three distinct clusters. All but nine items (A<sub>1</sub>, A<sub>6</sub>, A<sub>9</sub>, A<sub>13</sub>, A<sub>14</sub>, C<sub>2</sub>, D<sub>2</sub>, D<sub>4</sub>, D<sub>10</sub>) grouped together on three distinct factors. Each of the first three clusters was composed of items designed to measure a specific orientation. Ten of the fifteen items designed to measure the aggressive orientation loaded on the first factor, nine of the ten items designed to measure the compliant orientation loaded on the second factor, and seven of the ten items designed to measure the detached orientation loaded on the third factor. These factors were labeled accordingly. On item (D<sub>4</sub>) designed to measure the detached orientation and two items (A<sub>6</sub>, A<sub>9</sub>) designed to measure the aggressive orientation loaded on the fourth factor. This factor was not identified. The remaining six items did not load on any one factor. Three of these items (A<sub>1</sub>, A<sub>13</sub>, A<sub>14</sub>) were designed to measure aggressive orientation, two (D<sub>2</sub>, D<sub>10</sub>) to measure the detached orientation, and one (C<sub>2</sub>) to measure the compliant orientation.

A somewhat weaker argument could be made that item A<sub>14</sub> indicated the aggressive trait as expected since it loaded on the aggressive factor and the unknown factor. The same argument could also be made for items D<sub>2</sub> and D<sub>10</sub> in regard to the detached trait.

TABLE 1  
FACTOR LOADING MATRIX  
PRINCIPAL COMPONENTS -- VARIMAX ROTATION CONSTRAINED  
TO FACTORS WITH EIGENVALUES > 1

	1	2	3	4	5	6
A <sub>2</sub>	.517	.338	.130	-.278	.026	-.166
C <sub>2</sub>	.506	-.070	-.048	-.414	.144	.162
A <sub>7</sub>	.484	.279	.173	-.027	-.175	.040
C <sub>9</sub>	.404	-.382	-.025	-.028	-.296	-.183
C <sub>7</sub>	.387	-.284	-.217	.195	.157	.052
C <sub>3</sub>	.364	-.454	.079	-.268	-.080	-.027
C <sub>1</sub>	.356	-.398	.223	.036	.059	-.102
A <sub>10</sub>	.192	.574	.025	.145	-.029	.147
D <sub>7</sub>	-.182	.214	.493	-.071	-.083	.015
D <sub>5</sub>	-.173	.041	.485	.133	.089	.106
D <sub>9</sub>	-.393	.008	.477	-.202	.055	-.204
D <sub>10</sub>	-.161	.240	.459	.125	-.239	-.002
D <sub>2</sub>	-.017	.049	.424	.127	-.213	.201
D <sub>3</sub>	-.356	-.019	.338	-.075	.002	-.315
D <sub>8</sub>	-.109	.033	.239	-.229	-.350	-.006
A <sub>6</sub>	.249	.164	.052	.651	-.156	-.110
D <sub>4</sub>	.082	-.158	.162	.387	.160	.000
A <sub>8</sub>	.174	.296	.165	-.110	-.248	.491
A <sub>5</sub>	.536	.444	-.174	-.162	-.228	.054
A <sub>12</sub>	.372	.534	.005	.004	-.149	.105
A <sub>4</sub>	.335	.233	-.233	.134	-.207	-.047
A <sub>3</sub>	.326	.472	.221	-.140	.084	-.234
D <sub>6</sub>	-.115	.284	.267	-.394	.107	-.139
A <sub>14</sub>	.177	.274	.340	.233	.291	-.175
C <sub>10</sub>	.132	-.320	.237	.220	-.094	.303
C <sub>4</sub>	.354	-.330	.335	-.036	.234	-.028
C <sub>6</sub>	.349	-.523	.307	-.290	-.027	.113
C <sub>8</sub>	.293	-.508	.276	.053	-.222	.509
A <sub>13</sub>	.495	-.042	.139	.193	.359	-.166
A <sub>11</sub>	.482	.283	.093	.029	.341	.146
A <sub>15</sub>	.219	.080	.154	.133	.218	-.359
A <sub>9</sub>	-.069	-.087	.198	.271	-.123	.365
A <sub>1</sub>	-.039	.098	-.033	-.041	.459	.358
D <sub>1</sub>	-.283	.174	.127	-.181	.477	.418
C <sub>5</sub>	.197	-.497	.101	-.018	.010	.150
Eigen- value	3.537	3.406	2.254	1.725	1.572	1.436
% Var	10.11	9.73	6.44	4.93	4.49	4.10
Total Variance						39.8%

### Discussion

The principal component solution yielded a large number of factors, a small amount of explained variance, and the rotated solution did not approximate simple structure. On the other hand, the factor analysis solution yielded four factors and a large amount of explained common variance. The rotated factor solution approximated simple structure and the majority of items clustered according to design on three factors that appeared to represent Aggressive, Compliant, and Detached orientations.

The principal component results indicated that the majority of the variance in the instrument was either random or specific to the individual items. Since simple structure was not obtained, the random components appeared to be uncorrelated and large enough to preclude item clustering according to the design of the instrument. This supports the unreliability suspicion raised by the low internal consistency estimates.

TABLE 2

FACTOR LOADINGS MATRIX  
FACTOR ANALYSIS -- VARIMAX ROTATION CONSTRAINED  
TO FACTORS WITH EIGENVALUES > 1

## Conclusion

	A	C	D	?
A <sub>5</sub>	.609	-.084	-.243	-.204
A <sub>2</sub>	.591	.095	.010	-.174
A <sub>12</sub>	.558	.179	-.071	-.010
A <sub>3</sub>	.550	-.076	.136	-.027
A <sub>7</sub>	.515	.105	-.029	.031
A <sub>11</sub>	.504	.070	-.094	.037
A <sub>10</sub>	.456	-.300	-.016	.113
A <sub>8</sub>	.319	-.050	.114	-.040
A <sub>4</sub>	.292	-.080	-.290	.009
A <sub>15</sub>	.200	.085	-.001	.119
C <sub>6</sub>	.001	.681	.093	-.148
C <sub>8</sub>	-.053	.573	-.013	.136
C <sub>3</sub>	.004	.566	-.073	-.214
C <sub>1</sub>	.041	.502	-.061	.109
C <sub>4</sub>	.103	.489	.050	.075
C <sub>5</sub>	-.136	.445	-.081	.025
C <sub>9</sub>	.036	.437	-.227	-.015
C <sub>10</sub>	-.064	.319	.006	.211
C <sub>7</sub>	.032	.289	-.402	.060
D <sub>9</sub>	-.143	-.023	.546	.024
D <sub>7</sub>	.107	-.065	.463	.107
D <sub>3</sub>	-.168	-.043	.384	.053
D <sub>6</sub>	.163	-.118	.379	-.204
D <sub>5</sub>	-.005	.023	.353	.241
D <sub>1</sub>	-.047	-.193	.282	-.100
D <sub>8</sub>	.006	.016	.249	-.062
A <sub>6</sub>	.229	-.041	-.213	.495
C <sub>6</sub>	-.028	.154	-.038	.293
A <sub>9</sub>	-.069	.053	.070	.241
A <sub>13</sub>	.317	.294	-.157	.189
C <sub>2</sub>	.301	.319	-.135	-.331
A <sub>14</sub>	.319	-.013	.141	.272
D <sub>10</sub>	.113	-.114	.363	.310
D <sub>2</sub>	.086	.072	.248	.199
A <sub>1</sub>	.026	-.086	.017	-.046
Eigenvalue	2.790	2.654	1.808	1.130
% Common Variance	23.39	22.25	15.16	9.47
Total Common Variance 70.3%				

The factor analysis results indicated that when only a portion of the reliable (common) variance was analyzed the items generally behaved as expected. These findings, from both the principal components and factor analysis suggest that the domain of observables did not fit together. Although the instrument items related to each other in a consistent manner, these relationships appeared to have been eclipsed by instrument unreliability. This indicates that the instrument did not fulfill the second condition for construct validity.

The findings from this study are limited by the ungeneralizable sample. The instrument may perform differently on different target market segments. In fact Kernan (1971) reported different internal consistency estimates across five separate studies. The average estimate across these studies was .725 and the rank order of the three scales' reliability was generally the same as reported in this study. However, the possibility that the instrument may not operate as designed seems important enough to suggest that an analysis such as carried out in this research is appro-

priate for particular segments to which the instrument may be applied.

The evidence did not indicate that the instrument fulfills the second necessary condition for construct validity. It appears that the instrument has promise but more work is needed. The domain of observables should be enlarged so items can be added to each of the three scales. This is a generally accepted method of increasing scale reliability (Lord and Novick, 1968) that appears feasible since the instrument in its present form is easily completed in less than fifteen minutes. That is, respondent fatigue should not overcome reliability gains resulting from increased questionnaire length. As a first step, reliabilities should be increased until estimates are consistently above .80 so that correlation attenuations will not be a problem (Nunnally, 1967). After this is accomplished construct validation tests should again be undertaken.

The study also indicates that reliance on factor analysis in validation studies may produce misleading findings. Factor analysis, unlike principal components, does not consider instrument reliability which is a necessary condition for validity. The evidence suggests that utilization of both techniques may allow a researcher to ascertain whether reliability levels are low enough to invalidate an instrument.

The construction of a reliable and valid test instrument appears to be a tedious but necessary process. The validation work to date on the CAD scale has provided a foundation for continuation of the validation process. However, the instrument unreliability suggested by the findings in this study pose some problems. For example, Cohen and Golden (1972), in applying the instrument, report that differences in individual interpersonal orientations did not prove to be a significant factor in the acceptance of information from others. This failure to support hypothesized relationships may have been due to attenuations as a result of instrument unreliability. The lack of demonstrated instrument construct validity precludes what could otherwise be interpreted as disconfirmation of the theory since a valid instrument may have supported the relationships. The validation process should continue since it is doubtful that the instrument, in its present form, will allow a researcher to ascertain whether a failure to support or disconfirm hypotheses is the result of an inadequate theory or an inadequate methodology.

## References

- Becherer, Richard C., Jon F. Bibb, and Edward A. Riordan, "Spousal Perception of Household Purchasing Influence: A Multiperson-Multiscale Validation," Proceedings, Combined Conference, American Marketing Association, 1973, 289-292.
- Cohen, Joel B. "An Interpersonal Orientation to the Study of Consumer Behavior," Journal of Marketing Research, 4 (August 1967), 270-278.
- "Toward an Interpersonal Theory of Consumer Behavior," California Management Review, 10 (Spring 1968), 73-80.
- Cohen, Joel B. and Ellen Golden, "Informational Social Influence and Product Evaluation," Journal of Applied Psychology, 56 (February 1972), 54-59.

Cronbach, Lee J. "Coefficient Alpha and the Internal Structure of Tests," Psychometrika, 16 (September 1951), 297-334.

Davis, Harry S. "Measurement of Husband-Wife Influence in Consumer Purchase Decisions," Journal of Marketing Research, 8 (November 1974), 305-312.

Guttman, Louis, " 'Best Possible' Systematic Estimates of Communalities," Psychometrika, 21 (September 1956), 273-285.

Heeler, Roger M. and Michael L. Ray, "Measure Validation in Marketing," Journal of Marketing Research, (November 1974), 361-370.

Horney, Karen, Our Inner Conflicts. New York: W. W. Norton and Co., Inc., 1945.

Horton, Raymond L. "The Edwards Personality Preference Schedule and Consumer Personality Research," Journal of Marketing Research, 11 (August 1974), 335-337.

Kaiser, Henry F. "The Application of Electronic Computers to Factor Analysis," Educational and Psychological Measurement, 21 (Spring 1960), 141-151.

\_\_\_\_\_ "The Varimax Criterion for Analytic Rotation in Factor Analysis," Psychometrika, 23 (September 1958), 187-200.

Kassarjian, Harold H. "Personality and Consumer Behavior: A Review," Journal of Marketing Research, 8 (November 1971), 409-418.

Kernan, Jerome B. "The CAD Instrument in Behavioral Diagnosis," in D. M. Gardner, ed., Proceedings. The Association for Consumer Research, 1971, 307-312.

Kollat, David T., James E. Engel, and Roger D. Blackwell, "Current Problems in Consumer Behavior Research," Journal of Marketing Research, 7 (August 1970), 327-332.

Lord, Frederic M. and Melvin R. Novick, Statistical Theories of Mental Test Scores. Reading, Massachusetts: Addison-Wesley Publishing Company, 1968.

Nunnally, Jum C. Psychometric Theory. New York: McGraw-Hill, 1967.

Robertson, Thomas S. and Scott Ward, "Toward the Development of Consumer Behavior Theory," Proceedings. American Marketing Association, 1972, 57-64.

Thurstone, Louis L. Multiple Factor Analysis. Chicago: University of Chicago Press, 1947.

Appendix

C.A.D. QUESTIONNAIRE INSTRUCTIONS

Instructions: In this booklet you will find a number of incomplete statements followed by six blanks. These statements describe a variety of situations. There are no "right" or "wrong" answers. In fact, people's opinions regarding each statement seem to be quite different. The purpose of this survey will be served

best if you accurately report your feelings toward each statement. You may notice that many items are similar. Actually, no two items are exactly alike. Example:

	Extremely	Extremely
Asking a friend to	Undesirable	Desirable
loan you money is:	_____ X _____	_____

A check has been placed in the second blank. This means that the situation described is quite undesirable to the individual concerned.

These same six blank lines will be provided for each statement. Place a check mark on that blank which best expresses how desirable or undesirable the situation seems to you.

Name: \_\_\_\_\_ Age: \_\_\_\_\_  
 Sex: \_\_\_\_\_ Occupation: \_\_\_\_\_

Please turn the page and begin.

C.A.D. QUESTIONNAIRE ITEMS

- | Item Number*       |  |
|--------------------|--|
| (D <sub>1</sub> )  | 1. Being free of emotional ties with others is:  |
| (C <sub>1</sub> )  | 2. Giving comfort to those in need of friends is:  |
| (C <sub>2</sub> )  | 3. The knowledge that most people would be fond of me at all times would be:               |
| (A <sub>1</sub> )  | 4. To refuse to give in to others in an argument seems:                                    |
| (D <sub>2</sub> )  | 5. Enjoying a good movie by myself is:   |
| (D <sub>3</sub> )  | 6. For me to pay little attention to what others think of me seems:                        |
| (A <sub>2</sub> )  | 7. For me to be able to own an item before most of my friends are able to buy it would be: |
| (A <sub>3</sub> )  | 8. Knowing that others are somewhat envious of me is:                                      |
| (C <sub>3</sub> )  | 9. To feel that I like everyone I know would be:   |
| (D <sub>4</sub> )  | 10. To be able to work hard while others are elsewhere having fun is:                      |
| (A <sub>4</sub> )  | 11. Using pull to get ahead would be:  |
| (A <sub>5</sub> )  | 12. For me to have enough money or power to impress self-styled "bigshots" would be:       |
| (C <sub>4</sub> )  | 13. Basing my life on duty to others is:   |
| (A <sub>6</sub> )  | 14. To work under tension would be:  |
| (D <sub>5</sub> )  | 15. If I could live all alone in a cabin in the woods or mountains it would be:            |
| (A <sub>7</sub> )  | 16. Punishing those who insult my honor is:  |
| (C <sub>5</sub> )  | 17. To give aid to the poor and underprivileged is:  |
| (A <sub>8</sub> )  | 18. Standing in the way of people who are too sure of themselves is:                       |
| (D <sub>6</sub> )  | 19. Being free of social obligations is:   |
| (C <sub>6</sub> )  | 20. To have something good to say about everybody seems:                                   |
| (A <sub>9</sub> )  | 21. Telling a waiter when you have received inferior food is:                              |
| (D <sub>7</sub> )  | 22. Planning to get along without others is:   |
| (A <sub>10</sub> ) | 23. To be able to spot and exploit weakness in others is:                                  |
| (A <sub>11</sub> ) | 24. A strong desire to surpass others' achievements seems:                                 |
| (C <sub>7</sub> )  | 25. Sharing my personal feelings with others would be:                                     |
| (A <sub>12</sub> ) | 26. To have the ability to blame others for their mistakes is:                             |
| (D <sub>8</sub> )  | 27. For me to avoid situations where others can influence me would be:                     |
| (C <sub>8</sub> )  | 28. Wanting to repay others' thoughtless actions with friendship is:                       |
| (A <sub>13</sub> ) | 29. Having to compete with others for various  |



- rewards is:
- (D<sub>9</sub>) 30. If I knew that others paid very little attention to my affairs it would be:
  - (A<sub>14</sub>) 31. To defend my rights by force would be:
  - (C<sub>9</sub>) 32. Putting myself out to be considerate of others' feelings is:
  - (A<sub>15</sub>) 33. Correcting people who express an ignorant belief is:
  - (A<sub>10</sub>) 34. For me to work alone would be:
  - (C<sub>10</sub>) 35. To be fair to people who do things which I consider wrong seems:

\*Items are listed in order of appearance on the questionnaire. Item designations in parentheses were adopted to match items with their appropriate trait.

A STOCHASTIC INVENTORY DECISION MODEL  
AND THE HOLDING OF FAMILY WEALTH

Ivan F. Beutler, Iowa State University

Abstract

The decision process of managing family wealth is modeled. Stochastic inventory theory provided the underlying framework of the model by which family assets are viewed as inventories of wealth. The impact of family saving behavioral patterns upon asset management is also shown. Explicit functions predicting equilibrium cost and short term family asset holdings are derived.

Introduction

Optimal family wealth holdings and consumption patterns over the life cycle have long been of interest to the social scientist.<sup>1</sup> The consumption/wealth problem is a complex multiperiod decision process. This paper addresses only part of the overall problem by focusing on the holding of family wealth. This focus provides a unique understanding of consumer behavior different from that obtainable through a direct study of consumption. The focus of this paper can be understood by considering the portfolio problem which each family or household must solve.<sup>2</sup> Given that a family has a certain amount of wealth, that wealth must be held in some asset or combination of assets. The typical family holds a certain amount of their wealth in the form of currency and demand deposits. The remainder of their wealth is usually held in some combination of short and long term financial assets plus real assets, such as consumer durables and real estate.<sup>3</sup> The key question here is, what determines the combination, type and magnitude of assets held by the household? It is of interest to know what economic forces have a major bearing on the process, the quantitative magnitude of these forces and their role in shaping household portfolio behavior.

Background Literature

Keynes (1936) described three motives for holding cash; the transaction, precautionary and speculative motives. These motives have since served as the major underlying framework for most of the literature concerning the demand for money. Hicks (1935) recognized the potential importance of detail (such as transaction cost which he called "frictions") in an optimization-oriented approach to money demand.

Nearly two decades after Hicks stated that "frictions" were potentially important, Arrow, Harris and Marschak (1959), Whitin (1953) and others recognized "Inventory

Theory" as a potential tool and began to apply it in asset management theory.<sup>4</sup> Whitin's work in inventory management theory served as a suggestion to Baumol that it may have important application to the transaction demand for cash. The resulting article by Baumol (1952) is now a bench mark of the first inventory theoretic approach to cash management. However, the Baumol model employed a fairly simplified certainty construct. More recently others have expanded the Baumol certainty model to include various dimensions of household asset management.<sup>5</sup>

Miller and Orr (1966) introduced an inventory stochastic model for cash and asset management of the firm. The general flexibility and stochastic nature of the Miller-Orr model make it an important take-off from Baumol. Thus it was through an interest in the financial management of the firm at the micro level that the theory was further developed. Eppen and Fama (1969) have further expanded on the model within the context of the firm. However, to the writer's knowledge no serious attempt has been made to implement the structure into a theory of family asset management.<sup>6</sup>

The Stochastic Inventory Theoretic Decision Model

The family wealth or asset management decision process is viewed as an inventory problem. Each family has a certain amount of wealth which may be held in the form of cash, short term financial assets ("shorts"), long term financial assets ("longs") or real assets.<sup>7</sup> These four assets plus consumer credit sources are viewed as inventories subject to stochastic fluctuation due to cash inflow and outflow, consumption/saving decisions, and adjustments between the inventory assets.

Assumption One

A major assumption of the model is that each family has an asset manager who acts to minimize the expected cost

<sup>4</sup>Inventory theory goes back to the mid 1920's when development apparently began independently by several groups. Arrow (1958) makes reference to these beginnings.

<sup>5</sup>For example see Feige and Parkin (1971), Hamburger (1966), Sastry (1970), and Sprinkle (1969). Santomero (1974) sets forth a certainty model invoking the assumption of a uniform rate of change in the level of assets. From his model is then derived the intra-period average balance of currency, demand deposits, saving deposits and goods purchased as a function of the rates of return and transaction costs.

<sup>6</sup>Orr (1971) Chapter 7, makes several innovative suggestions concerning household demand for money and further suggests that the basic model developed by himself and Miller might be modified for the household setting.

<sup>7</sup>Cash is defined here to include currency and demand deposits.

<sup>1</sup>Fisher (1930) provided micro foundations for much of modern macro theory of aggregate saving, wealth and consumption. The focus here is at the micro level for individual households.

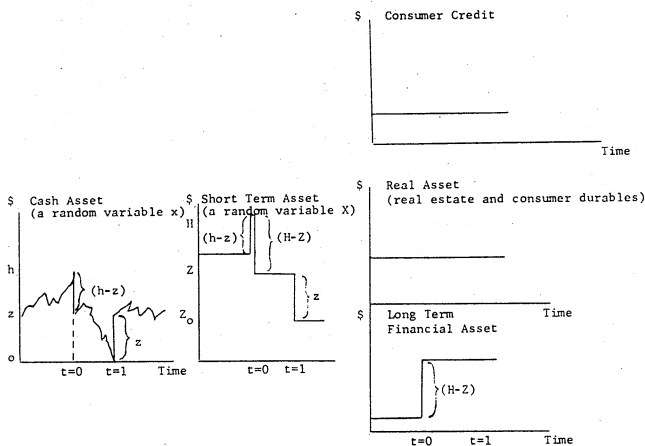
<sup>2</sup>The solution may range from following an elaborate plan to that of simple default behavior.

<sup>3</sup>Such factors as human wealth are obviously not considered in the management of wealth as used here.

of managing household assets by following certain policy rules of behavior.

**Rule 1. (h-z) Policy Rule.** The family asset manager is assumed to follow an (h-z) policy rule. This management policy is illustrated in Figure 1. Cash balances are allowed to fluctuate freely until either an upper bound h or a lower bound zero is reached. When cash accumulates to the upper bound h, the asset manager effects a portfolio transfer of (h-z) dollars into the short term asset account. When cash balances fall to the zero boundary, z dollars are transferred from the short term asset ("shorts") into cash, restoring the cash balance to level z. Likewise, when the upper bound H is reached in the shorts account, (H-Z) dollars are transferred into long term assets ("longs") or the purchase of real assets.<sup>8</sup> When the value of short term assets is reduced to the lower bound  $Z_0$ , short term assets are restored to a level Z either by the use of credit, or by the liquidation of long term financial assets or real assets. The source used depends on three factors: 1) the transaction cost of liquidating either "longs" or real assets versus using credit, 2) the cost of credit compared to the differential return on longs versus real assets, and 3) the cost of credit compared to the differential return on long and short term assets.<sup>9</sup> The short term asset lower bound  $Z_0$ , will vary from family to family and depends on the household's present-future orientation.

FIGURE 1. An Asset Management Policy Rule (h-z).



In short, the family asset manager is assumed to respond to cash balance x, short term asset X, and long term asset and real assets Y, situations as follows:

<sup>8</sup> Consumer credit debt repayment is assumed to take place via installment payments. Thus credit repayments are made out of the cash account.

<sup>9</sup> The transaction cost of liquidating "longs" will depend on the dollar amount of funds transferred. The transaction cost will be infinite if a positive level of long term assets does not exist. The cost of credit will depend on the risk class into which the particular family fits as viewed by creditors. It should be noted that the return on real assets such as consumer durables may be an implicit return as well as financial.

Situation	Asset Management Response
$x = 0, Z_0 \leq X < z,$ $Y \geq z + Z - X$	Increase both cash and shorts via-sell $z+Z-X$ of longs or real assets, or obtain $z+Z-X$ of credit <sup>10</sup>
$x = 0, Z_0 \leq X < z,$ $Y < z + Z - X$	Increase both cash and shorts via-obtain $z+Z-X$ of credit
$x = 0, Z_0 + z \leq X$	Increase cash only via-sell z of shorts
$0 < x < h$	Do nothing
$x = h, X < H - (h - z)$	Reduce cash only via-buy h-z of shorts
$x = h, X \geq H - (h - z)$	Reduce both cash and shorts via-buy h-z + X-Z of longs or real assets and sell X-Z of shorts

**Rule 2. (c-s) Policy Rule.** The (c-s) policy rule takes into account two types of family saving behavior patterns not accounted for as yet in the model. These saving patterns may affect cash balances substantially. Manning (1960) has reported three distinct family saving practices. The first practice is modeled above by the (h-z) policy rule. 1) Residual Saving - this is the traditional view of saving where money left over after personal consumption expenditures have been taken from disposable income is considered to be a residual balance. When these residual balances have accumulated to a certain level and are transferred into an interest bearing asset it is considered, by definition, to be residual saving. 2) Fixed Commitment Saving - fulfilling a commitment to save or set aside funds before consumption demands are met. 3) Special Receipt Saving - Setting aside special receipts or income (such as a tax refund).

Committed saving and the saving of special receipts require separate treatment in the model. These two practices give rise to the (c-s) management policy (committed and special receipt saving). All (c-s) managed funds which are used for the purchase of short term assets, by-pass the cash balance account and go directly for the purchase of "shorts". This is so by definition of committed and special receipt saving. Furthermore, for the purpose of this model, it is assumed that, when these (c-s) short term assets are liquidated, purchase synchronization is perfect.<sup>11</sup> Thus (c-s) saving effects both the mean and variance of short term asset holdings but, the cash balance stationary density function remains unaltered.

#### Assumption Two

A second assumption of the model is that the level of each asset inventory is stochastic in nature such that it can be treated as a random variable.

<sup>10</sup> A separate study is necessary to determine the relationship between the liquidation of long term assets, real assets and the use of credit.

<sup>11</sup> Perfect synchronization assumes that the purchase of goods and services from (c-s) short term assets is coincidental with (c-s) liquidation. Thus the assumption is, that of all short term assets liquidated over time, a dollar amount equal to (c-s) saving goes directly for the purchase of goods and services.

Cash Balance Asset, a Random Variable-x. To keep the model manageable, a simple Bernoulli transition rule is assumed to describe changes in the cash balance random variable  $x$ . Whenever a state exists, the Bernoulli assumption is that during the succeeding state of the process, cash balances will increase with probability  $p$ , decrease with probability  $q$ , and that  $p + q = 1$ . Thus the state-to-state transitions can be defined as

$$P_{t+1}(x) = qP_t(x+1) + pP_t(x-1) \quad 0 < x < h, x \neq z$$

$$P_{t+1}(z) = q[P_t(x+1) + P_t(z+1)] + pP_t(z-1) + pP_t(z-1), x=z$$

$$P_{t+1}(0) = P_{t+1}(h) = 0$$

$$p + q = 1$$

$P_{t+1}(x)$  = the probability that cash balances will be at the level  $(x)$  at time  $t+1$ .

By invoking the steady state condition, the limit as  $t$  approaches infinity, the transition rule can be

$$\lim_{t \rightarrow \infty} P_{t+1}(x) = P(x)$$

characterized by the following difference equations.

$$P(x) = pP(x-1) + qP(x+1) \quad 0 < x < h \quad x \neq z$$

$$P(z) = p[P(z-1) + P(h-1)] + q[P(z+1) + P(1)] \quad x=z$$

$$P(0) = P(h) = 0$$

$$\sum_0^h P(x) = 1$$

$$x = 0$$

The equation listed above is a homogeneous difference equation which can be readily recognized by normalization,

$$P_{x+1} - 1/qP_x + p/qP_{x-1} = 0.$$

The auxiliary equation is of the form

$$m^2 - (1+p/q)m + p = 0.$$

Letting  $p/q = r$ , the roots of the auxiliary equation can be easily seen.

$$(m_1 = 1)(m_2 = r)$$

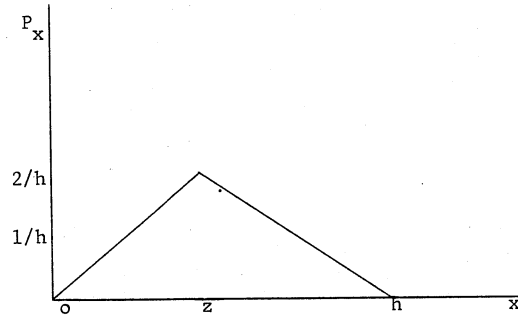
Hence the general solution form of the homogeneous equation can be solved to obtain the stationary density function describing the random variable  $x$ .<sup>12</sup> The resulting function (15) is a triangular, probability density function illustrated

$$P_x = \begin{cases} zx/zh & 0 \leq x \leq z \\ z(h-x)/h(h-z) & z \leq x \leq h \end{cases} \quad (1)$$

<sup>12</sup>For derivational detail see Beutler (1974).

in Figure 2.<sup>13</sup> The nature of the random variable  $x$  (cash balances) is described by the density function, which has, as its parameters, the policy control limits  $z$  and  $h$ . This density function contains all the information about  $x$  which is necessary to derive the expected asset management cost function.

FIGURE 2. Triangular Stationary Probability Density Function - Generated by a Symmetric Bernoulli Markov Process.



Short Term Assets, a Random Variable-(X). It is also of interest to know the stochastic nature of the random variable  $X$ . The level of short term asset holdings has its origin from two sources. The first source is the  $(h-z)$  policy rule. An accumulation and liquidation of "shorts" behaving in accord with this rule creates a component of the short term asset which behaves as a random variable, denoted  $u$ . The  $(c-s)$  policy rule creates a second component of short term asset holdings which also behaves as a random variable, denoted  $s$ .

Short term asset behavior is thus described as the sum of two random variables, so that  $X = s + u$ . Since  $s$  is exogenous to the model,  $s$  and  $u$  are taken to be stochastically independent. The stochastic nature of the random variables  $u$  and  $s$ , can be formulated as separate asymmetric Bernoulli processes. These may then be pooled to determine the stochastic nature of  $X$ . However, the error introduced into the mean balance estimator, by using a single-step symmetric Bernoulli estimate  $(H+Z)/3$  to represent the complex pooled process, has been shown to be remarkably small and virtually independent of the values  $H$  and  $Z$ .<sup>14</sup> Thus, the short term asset mean value of the pooled process  $X = u + s$ , may be estimated by equation (2).

$$E(X) = (H+Z)/3 \quad (2)$$

<sup>13</sup>By the assumption of nonpersistent drift ( $r = p/q = 1$ ), the density function given in (1) is considerably simplified from the more general form.

$$P_x = \begin{cases} \frac{1}{z} + \frac{1-r^{z-h}}{(1-r^z)(h-z)} (1-r^x) & 0 \leq x \leq z \\ \frac{1-r^z}{z(1-r^{z-h}) + (h-z)(1-r^z)} (1-r^{x-h}) & z \leq x \leq h \end{cases}$$

<sup>14</sup>Feller (1966) and Orr (1974) have shown that small error results when complex pooled process mean estimates are replaced by simpler approximation estimates from a single process. Furthermore, Miller and Orr (1968) have shown  $(H+Z)/3$  to be a reasonably good estimate of  $\bar{x}$  regardless of the level at which  $h$  and  $z$ , the cash balance policy parameters, are set and regardless of the level of  $H$  and  $Z$ .

Since  $u$  and  $s$  are taken to be independent random variables, the variance of  $X$ ,  $\sigma_X^2$ , is the sum of the two variances  $\sigma_X^2 = \sigma_s^2 + \sigma_u^2$ . First consider the variance of  $u$ ,  $\sigma_u^2$ . It is given by the following equation.<sup>15</sup>

$$\sigma_u^2 = \sigma^2/z(h-z) \left[ \frac{(h-z)(z/h)}{(z/h)(h-z)} \right]^2 z/h + \left[ \frac{(z/h)(h-z)}{(h-z)/h} \right]^2 (h-z)/h = \sigma^2 \quad (3)$$

Thus the variance of the periodic movement between "shorts" and cash is  $\sigma^2$ . That the variance of this short term asset component  $u$ , and the cash balance asset  $x$ , both equal  $\sigma^2$  is no trivial matter. Since the short term asset variance does not depend directly upon  $z$  and  $h$ , the derivation of optimal policy controls is much less difficult. This will become clear when the probability of a portfolio adjustment between "shorts" and "longs" is considered in the following section.

A second contribution to deviation in short term asset holdings is the variance of the random variable  $s$ . The variance of this exogenous variable is denoted  $\sigma_{c-s}^2$ . Thus, the total variance of the short term asset is

$$\sigma_X^2 = \sigma^2 + \sigma_{c-s}^2 \quad (4)$$

#### Assumption Three

So far two aspects of the model have been established; 1) each family asset manager follows an  $(h-z)$  and  $(c-s)$  policy rule of behavior, and 2) the stochastic nature of the asset inventories is characterized as a symmetric Bernoulli process. Now the additional assumption is made that asset managers choose the optimal policy controls  $z$ ,  $h$ ;  $Z$  and  $H$  so as to minimize the expected cost of asset management. This requires that an expected cost function be formulated and minimized with respect to  $z$ ,  $h$ ,  $Z$ , and  $H$  as expressed in general form in the following equation.

$$\text{Minimize } E(c) = T_C P(T_C) + i_l E(x) + \alpha G(x) + T_S P(T_S) + (i_l - i_s) E(X) \quad (5)$$

where:

$E(c)$  = The expected cost of asset management.

$P(T_C)$  = The probability of a transfer between short term and cash balance assets.

$E(x)$  = The expected level of cash balances.

$G(x)$  = The cost function representing the price which households place on holding wealth in a liquid form. A form more subject to impulsive, short run purchases which may be suboptimal in the longer run.

$P(T_S)$  = The probability of a transfer between short term assets and credit or long term assets.

$E(X)$  = The expected level of short term assets.

<sup>15</sup>In words the variance of  $u$  reads  $\sigma^2 = E(u-\bar{u})^2 = [\text{expected increase in "shorts"}]^2 (\text{probability of an increase}) + [\text{expected decrease in "shorts"}]^2 (\text{probability of a decrease})$ .

$T_C$  = A lump sum transfer cost associated with transfers between the cash balance and short term assets. This cost includes the value of management time and other indirect costs, as well as, direct outlays, such as gasoline for a trip to the bank, postage and brokerage fees.

$i_l$  = The earning rate on long term assets which represents the interest opportunity cost of holding cash.

$i_s$  = The earning rate on short term assets.

$\alpha$  = The rate or price which households attach to the holding of liquid wealth.

$T_S$  = A lump sum transfer cost associated with transfers between short term assets and credit, long term assets, or real assets.

From the cost function  $T_C$ ,  $i_l$ ,  $\alpha$ ,  $T_S$  and  $(i_l - i_s)$  are all fixed price parameters which do not vary with respect to the level of the policy controls  $z$ ,  $h$ ,  $Z$  and  $H$ .  $P(T_C)$ ,  $E(x)$ ,  $G(x)$ ,  $P(T_S)$ , and  $E(X)$  are each stochastic variables, the value of which depends on the level at which the policy controls are set.

The probability of a transfer from the cash account into the short term asset account is denoted  $P(T_{C-S})$ . Likewise, a transfer from "shorts" into the cash account is denoted  $P(T_{S-C})$ . The probability of a  $T_{C-S}$  transfer during any given time period is equal to the probability of a state  $h-1$  occurrence, times the probability of a transfer, given the state  $h-1$ .<sup>16</sup> Similar reasoning holds for the probability of a  $T_{S-C}$  transfer. The total probability of a transfer, irrespective of the direction, is equal to  $P(T_{C-S})$  plus  $P(T_{S-C})$  as shown in equation (6).

$$P(T_{C-S}) = \sigma^2/h(h-z) \quad (6.a)$$

$$P(T_{S-C}) = \sigma^2/zh \quad (6.b)$$

$$P(T_C) = P(T_{C-S}) + P(T_{S-C}) = \sigma^2/z(h-z) \quad (6.c)$$

The average cash balance or mean value of  $x$  is given by equation (7).

$$E(x) = \int_0^h x P_x dx = (h+z)/3. \quad (7)$$

The first two components of the expected cost function, equation (5), are seen to be readily stated in explicit form, equation (6) and (7), once the stationary distribution function of equation (1) is known.

<sup>16</sup>The probability of a state  $h-1$  occurrence during a given time period is equal to the number of transactions which occur in the cash account during that period, times  $P_{h-1}$ . The unit of change in the simple Bernoulli process is  $\sigma^2$ . For detail on this point see Wold (1946). The probability of a transfer, given the occurrence of the state  $h-1$ , equals  $1/2$ . Thus we have

$$P(T_{C-S}) = \sigma^2 [2/h(h-z)] (1/2) = \sigma^2/h(h-z).$$

<sup>17</sup> $E(x) = \int_0^z x P_x dx = \int_0^z 2x^2/zh dx + \int_z^h 2x(h-x)/h(h-z) dx = (h+z)/3.$

Procedure for derivation of the third cost component  $G(x)$ , is not as clearly defined as were the first two. There is no research, to the author's knowledge, which suggests the explicit form the  $G(x)$  cost function should take. Thus for lack of a better form,  $G(x)$  is assumed here to be a simple positive function of the expected level of cash balances,

$$G(x) = E(x) = (h+z)/3. \quad (8)$$

The last two cost components of equation (5) are related to the holding of short term assets. Since the variance of movement in the shorts account is  $\sigma_x^2 = \sigma^2 + \sigma_{c-s}^2$ , given by equation (4), the expected first passage time for the shorts balance from  $Z$  to either 0 or  $H$  is  $Z(H-Z)/(\sigma^2 + \sigma_{c-s}^2)$ . The steady-state probability of a first passage, which signals a purchase or sale of longs, is the inverse of the first passage time  $(\sigma^2 + \sigma_{c-s}^2)/Z(H-Z)$ . Thus the explicit function of expected costs, resulting from a portfolio transfer between "shorts" and "longs", real assets or credit, is

$$P(T_s) = \sigma_x^2/Z(H-Z). \quad (9)$$

The expected level of short term asset holdings is given by equation (2) as  $E(X) = (H+Z)/3$ . However, this expression was derived without taking into account the reserve level of short term asset  $Z_0$  held by families.  $Z_0$  will vary from household to household, but is assumed to be a constant for any given family. The fifth and final component of the expected cost function is thus

$$E(X) = (H+Z+Z_0)/3. \quad (10)$$

The expected cost minimization problem stated explicitly in terms of the policy controls  $h$ ,  $z$ ,  $H$  and  $Z$  follows directly from equation (5) through (10).

The optimization object (as stated by equation (5)) is to set the policy parameters at an optimal level, denoted  $z^*$ ,  $h^*$ ,  $Z^*$  and  $H^*$ , such that the expected cost of asset management is minimized.

$$\begin{aligned} \text{Minimize } E(c) &= \frac{i_l(h+z)}{3} + \frac{T_c \sigma^2}{z(h-z)} + \frac{\alpha(h+z)}{3} + \\ \text{w.r.t. } z, h, Z, H & \frac{(i_l - i_s)(H+Z+3Z_0)}{3} + \frac{T_s(\sigma^2 + \sigma_{c-s}^2)}{Z(H-Z)} \end{aligned} \quad (11)$$

The necessary conditions for a minimum are given by equation (12).

$$\frac{\partial E(c)}{z} = \frac{i_l + \alpha}{3} - \frac{T_c \sigma^2 (h-2z)}{z^2(z-h)^2} = 0 \quad (12.a)$$

<sup>18</sup>Families which do not have a residual saving motive of the  $G(x)$  type, will have a price vector equal to zero. In this event the  $\alpha G(x)$  cost component will fall out of the expected cost function.

<sup>19</sup>Equation (9) may appear ordinary, but a very significant point should be noted. The probability of a transfer between "longs" and "shorts" is seen to depend only on the level of  $Z$ ,  $H$  and  $\frac{Z}{X}$ . It does not depend directly on  $h$  or  $z$ . The minimization problem would be much more difficult if  $h$  and  $z$  came into the cost expression (9).

$$\frac{\partial E(c)}{h} = \frac{i_l + \alpha}{3} - \frac{T_c \sigma^2 z}{z^2(z-h)^2} = 0 \quad (12.b)$$

$$\frac{\partial E(c)}{Z} = \frac{(i_l - i_s)}{3} - \frac{T_c(\sigma^2 + \sigma_{c-s}^2)(H-2Z)}{Z^2(H-Z)^2} = 0 \quad (12.c)$$

$$\frac{\partial E(c)}{H} = \frac{(i_l - i_s)}{3} - \frac{T_c(\sigma^2 + \sigma_{c-s}^2)Z}{Z^2(H-Z)^2} = 0 \quad (12.d)$$

The above first order conditions are set equal to zero and solved simultaneously for the optimal level of each policy control variable.<sup>20</sup>

$$z^* = \left( \frac{3T_c \sigma^2}{4(i_l + \alpha)} \right)^{1/3} \quad h^* = 3z \quad (13)$$

$$Z^* = \left( \frac{3T_s(\sigma^2 + \sigma_{c-s}^2)}{4(i_l - i_s)} \right)^{1/3} \quad H^* = 3Z \quad (14)$$

In terms of the optimal policy control limits, the expected cash balance is equal to  $(h^* + z^*)/3$ . Substituting into this expression the value of  $h^*$  and  $z^*$  from equation (13), gives the average cash balance  $\bar{x}$ .

$$\bar{x} = 4/3 \left( \frac{3T_c \sigma^2}{4(i_l + \alpha)} \right)^{1/3} = 4/3z \quad (15)$$

The average cash balance  $\bar{x}$ , is a measure of average family cash holdings. Likewise, the average holding of short term financial assets is given by  $\bar{X}$ .<sup>21</sup>

$$\bar{X} = 4/3 \left[ \left( \frac{3T_s(\sigma^2 + \sigma_{c-s}^2)}{4(i_l - i_s)} \right)^{1/3} + \frac{3Z_0}{4} \right] = \frac{4Z + 3Z_0}{3} \quad (16)$$

Although the mechanics of the above optimization problem get involved, the essence of the problem is straight forward. By minimizing the expected cost function, the model-predicted, optimal policy controls are given by equations (13) and (14). The model predictions of average asset holdings, are also given by equations (15) and (16). Thus the basic three assumptions of the model culminate in equations (13) through (16) to provide; 1) the optimal level at which policy controls would be set to minimize management cost, and

<sup>20</sup>For details of this solution see Beutler (1974). Let  $h=kz$  and  $H=KZ$ . The second order conditions for a minimum solution are met so long as  $k > 2.7$  and  $K > 2.7$ . With these constants greater than 2.7, the Hessian Matrix formed from equation (13) is positive definite. The positive definite Hessian Matrix implies that the expected cost function is strictly convex and thus the first order conditions lead to a global minimum solution.

<sup>21</sup>These results are equivalent to those of Miller and Orr (1966) with the exception of  $Z_0$  in equation (16) and the variance of short term assets equal to  $\sigma^2 + \sigma_{c-s}^2$  instead of simply  $\sigma^2$ .

2) an explicit expression of family asset holdings in equilibrium given basic information about the economic and family environment.

#### Model Implications

The optimal policy control limits  $z^*$  and  $h^*$ , are seen from equation (13), to vary directly with the cash balance variance ( $\sigma^2$ ). On a priori grounds, it seems correct that a family with larger (rather than smaller) cash inventory variation will hold larger (rather than smaller) cash balances. Thus, a family which has large variation in cash balance holdings over time, will optimally maintain a high  $z$  and  $h$  level. This result is intuitively appealing and provides a facet of the theoretic model which can be tested against the data. Note that  $\sigma^2$  serves to generalize the model by providing a mechanism whereby a wide range of management behavior is explained through a single parameter. Family differences such as stages in the life cycle, income level, and expenditure patterns all have a likely impact on  $\sigma^2$  which in turn, accounts for a wide range of management behavior.<sup>22</sup> For example, households which synchronize cash inflow and outflow, by the use of credit and other payments, will reduce cash balance variations. In fact,  $\sigma^2$  is a measure of inflow-outflow synchronization.

The mean cash balance  $x$ , and the optimal policy control limits ( $z^*$  and  $h^*$ ), are seen to be an increasing function of the transaction cost ( $T$ ) and a decreasing function of the opportunity interest rate ( $i$ ) and the management price of holding cash ( $\omega$ ). When a family accumulates short term assets with a residual saving motive coupled with the interest earning incentive,  $\alpha$  will be positive. A positive  $\alpha$  lowers the optimal policy controls ( $z^*$  and  $h^*$ ). This creates more frequent transfers between short term assets and the cash accounts, generating greater residual saving. If no residual saving motive exists, beyond the interest incentive,  $\alpha$  will equal zero and average cash holdings will be independent of this motive. The direction of impact which these variables ( $T$ ,  $i$  and  $\omega$ ), have on cash balances is in accord with economic theory.

#### Discussion

This paper represents a focus on consumer behavior which is not a direct act of consumption but rather a closely related process of asset management for the purpose of postponed consumption. The unique contribution of the stochastic inventory model set forth here, is that the family wealth and asset management decision process is characterized in a somewhat complex, but manageable, world including uncertainty. In contrast, the Baumol model and extensions thereof, assume a simple world of certainty where cash inflows take the form of income at a constant rate placed directly into an interest bearing asset. The result being that portfolio transfer costs and the interest opportunity cost of holding cash are the only variables perceived as relevant in explaining the decision process. On the other hand, the asset management of family wealth as modeled here, more completely exposes the actual environment and decision process by including a form of uncertainty, family saving behaviors, and family propensity to manage (cash/saving).

<sup>22</sup>For the purpose of expedition,  $\alpha$  is assumed to equal zero here.

#### References

- Kenneth J. Arrow, "Historical Background" in Arrow, Karlin and Scarf (ed.), Studies in the Mathematical Theory of Inventory and Production (Stanford: Stanford University Press, 1958).
- Kenneth J. Arrow, Theodore Harris, and Jacob Marschak, "Optimal Inventory Policy," Econometrica 19, 1951, 250-272.
- Robert J. Barrow and Anthony M. Santomero, "Household Money Holdings and the Demand Deposit Rate," Journal of Money, Credit and Banking 4(2), 1972, 397-413.
- William J. Baumol, "The Transactions Demand for Cash: An Inventory Theoretic Approach," Quarterly Journal of Economics 66, 1958, 545-556.
- Ivan F. Beutler, Family Cash Management-A Stochastic Inventory Theoretic Model, Ph.D. Dissertation, 1974.
- Gary D. Eppen and Eugene F. Fama, "Cash Balance and Simple Dynamic Portfolio Problems with Proportional Costs," International Economic Review 10, 1969, 119-133.
- William Feller, An Introduction to Probability Theory and Its Applications (New York: John Wiley & Sons. Vol. I (discrete sample spaces) 1951, Vol. II (continuous sample spaces), 1966).
- Irving Fisher, The Theory of Interest (New York: Augustus M. Kelly, Publishers, 1965 - Reprinted from the 1930 edition).
- J. M. Keynes, The General Theory of Employment, Interest and Money (New York: Harcourt-Brace & Co., 1936).
- Sarah L. Manning and Ethel L. Vatter, "Financial Management Practices of Families with Steady or Fluctuating Incomes," Cornell University Agricultural Experiment Station, Bulletin 997, New York State College of Home Economics, Ithaca, New York, 1964.
- Merton H. Miller and Daniel Orr, "The Demand for Money by Firms: Extensions of Analytic Results," Journal of Finance 23, 1968, 735-759.
- Merton H. Miller and Daniel Orr, "A Model of the Demand for Money by Firms," Quarterly Journal of Economics 80, 1966, 413-435.
- Daniel Orr, Cash Management and the Demand for Money (New York: Praeger, 1970).
- Anthony M. Santomero, "A Model of the Demand for Money by Households," Journal of Finance 42, 1974, 89-101.
- Rama A. S. Sastry, "The Effect of Credit on Transactions Demand for Cash," Journal of Finance 25, 1970, 777-781.
- Case M. Sprenkle, "The Uselessness of Transactions Demand Models," Journal of Finance 24, 1969, 835-847.

## A NORMATIVE DEFICIT MODEL OF CONSUMER BEHAVIOR

Earl W. Morris, Iowa State University  
Mary Winter, Iowa State University  
Ivan F. Beutler, Iowa State University

### Abstract

A model is developed in which consumer behavior is seen as a response to changes in satisfaction as a result of normative deficits in the current level and quality of consumption. Normative deficits arise from the difference between consumption as prescribed by cultural and personal norms and achieved consumption.

### Introduction

This paper outlines the development of a normative deficit model of consumer behavior. The model integrates both the socialpsychological and sociocultural approaches to the study of consumer behavior. The primary focus is on the cultural sources of the motivation to augment, improve, or otherwise alter the current supply of one or more of the attributes of various consumer goods. The present model elaborates upon two sets of concepts included in many of the other integrative models of consumer behavior: motives and their source in cultural norms.

This paper derives from a view of consumption as an adjustment mechanism to maintain a balance between needs, defined in terms of cultural norms (Tarde, 1903:44; Parsons and Shils, 1954:9-10), and achieved levels of consumption. The family's current level of consumption is evaluated by comparison with two sets of standards: (1) strictly cultural norms and (2) family level norms that more or less closely approximate the cultural norms. A normative deficit is based on the difference between the current supply (at the household level) of a good and the cultural standards and/or personal standards for that good.

Consumption, therefore, is viewed as an ongoing evaluation process conducted more or less continuously and more or less consciously by all families. The attention of this paper therefore is not focused upon consumption for subsistence. This approach seems sound since the model is limited to the United States (although it clearly could be adapted to other societies) and few Americans consume at or near the subsistence level.

The basic orientation of the paper arises not from consumer research in the marketing tradition, nor directly from traditional economic theory, although it is consistent with the latter. In fact, it can serve as an elaboration of the concept of nonsatiation and the process by which utility is maximized.

Rather, the roots of the model are found in the developmental approach to the study of the family (Hill and Hansen, 1960; Hill and Rodgers, 1964; Rodgers, 1962; Rodgers, 1964). In particular, the maintenance of a low total normative deficit can be seen as one of the dominant developmental tasks for the family. Further, changes in the family composition by age and sex over the history of the family by themselves bring about changes in the total deficit because the changes in the family can require that new standards be applied. Different standards would apply for a family with teenage children than a family with infants or no children.

Thus, a task for the family is to maintain the normative deficits at a low level by adjusting to the progression of standards that apply over the life cycle.

The model is a generalization of an earlier work on housing choice and adjustment (Morris and Winter, 1975). Normative housing deficits were found to be related to residential satisfaction (Morris, Crull and Winter, 1975), residential mobility (Gladhart, 1973; Morris and Gladhart, 1975), the propensity to move (Morris, Crull and Winter, 1975), residential alteration and improvement (Bross, 1975) and indirectly to the use of contraception to avoid residential crowding (Bresler, 1975).

### Other Models of Consumer Behavior: A Brief Overview

A number of integrative models of consumer behavior have appeared in the literature. One of the earliest of these was by Howard (1963), followed by Nicosia (1966). Engel, Kollat and Blackwell (1968) expanded greatly upon the earlier models and developed a detailed diagrammatic attitudinal model of the consumer decision process.

According to Kollat, Blackwell and Engel (1972), none of the previous integrative models had greatly influenced consumer research, including, apparently, Engel, Kollat and Blackwell (1968). One of the main reasons for the relative disuse of the integrative models has been the lack of clarity as to how most of the concepts in the models might be operationalized. In addition, it has not always been clear if a decision net/flow chart type of model was being presented, or whether a causal model was intended.

Another limitation of consumer attitude research has been its narrow socialpsychological focus. Sociocultural factors, when included, were seldom elaborated (see, for example, Engel, Kollat and Blackwell, 1968). Rising interest in the role of social and cultural influences on consumer behavior are seen in Ward's (1974) review of research on consumer socialization and the frequent, albeit cursory, mention in numerous research reports of the influence of cultural norms (see, for example, Wilson, Mathews and Harvey, 1975).

A great deal of recent research has been focused on the testing of brand choice attitudes and behavior, based on models developed by Rosenberg (1956) and Fishbein (1963). Nakanishi and Bettman (1974), Olshavsky and Summers (1974), Bettman, Capon and Lutz (1975) and Wilson, Mathews and Harvey (1975) all begin with a presentation of the Fishbein model (adapted to brand choice) and then proceed to test some portion of it or attempt to improve its applicability by introducing additional factors.



The models, as tested, have typically taken the form:

$$A_j = w_i s_i \quad (1)$$

where:  $A_j$  is the overall rating of brand  $j$   
 $w_i$  is the belief that brand  $j$  possesses attribute  $i$   
 $s_i$  is the attitude toward attribute  $i$  of brand  $j$

In some cases  $w_i$  is defined as the importance of attribute  $i$ . Typically, the importance weights are estimated by regression coefficients but some studies test the model using belief or importance ratings provided by the respondent. Both the regression form and the subjective weighting form of the model suffer from extreme oversimplicity. It is doubtful that even an act as simple as choosing a brand of toothpaste can be satisfactorily explained or predicted by the use of attitudes and beliefs alone.

Even less appropriate is the use of simple attitude-belief models for the explanation of more global consumer decisions such as whether to purchase a given good at all, as opposed to choosing a specific brand of that good. Moreover, the emphasis on the attitudes toward the attributes of specific brands is misplaced since the attention of the consumer is first focused on the adequacy of the current supply of the product.

#### Normative Deficits and Consumer Behavior

The basic concept included in our model that appears to be absent from previous consumer research and integrative models of consumer behavior is the notion of normative deficits. With reference to housing, for example, the number of bedrooms needed depends upon cultural standards for bedroom sharing. Subtracting the number of bedrooms needed from the number in the current residence gives a variable that may range from negative numbers indicating a deficit to positive numbers indicating a surplus of bedrooms. A summation of a series of normative deficits is seen as an intervening variable that explains the relationship between a number of exogenous variables and general satisfaction (see Figure 1).

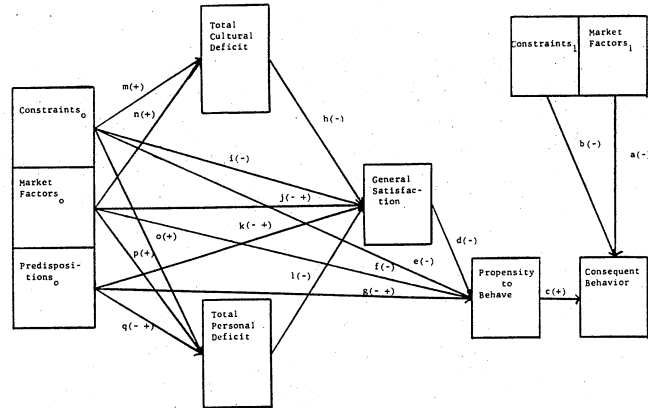
The motivation for consumption behavior arises from a dropping of the overall deficit level below tolerable limits. The consumer is motivated to make a new purchase when the current supply is deficient in quantity or quality as judged by the consumer and/or the society. This point is stressed in order to avoid the psychologism of assuming that purchases result solely from the attitudes of the individual.

Examples of normative deficits can be drawn from consumer behavior in regard to housing. Suppose cultural standards require three bedrooms in a dwelling for a family with husband and wife and two teenaged children, one a boy and one a girl. If such a family lives in a dwelling with only two bedrooms, that family would experience a cultural normative deficit for the number of bedrooms. Two families of equal composition may experience the same cultural norms and corresponding normative deficits, but have different personal normative deficits. One family might experience a personal deficit if their residence did not provide a garden space. There are, of course, many additional deficits that could appear, depending upon the attributes of their current dwelling.

The deficit type of variable would be unnecessary if all individuals applied a common standard to their situation. If that were the case, a description of, as examples, the number of rooms in their dwelling,

the number of floors, the amount of yard space, etc., would be sufficient. Such is not the case, since many individuals apply unique standards to themselves. Further, common standards may apply to some situations and not to others. An example is the standard for the number of rooms prescribed for a dwelling which is contingent upon household size and composition.

FIGURE 1. Causal Diagram of Analytical Model



The deficits are important as they are seen as producing motivations to "do something" about, for example, housing, such as move to a different dwelling. Normative deficits arise from the attributes of the existing dwelling, or current life conditions, and not from the potential attributes of the new dwelling. Thus, a key difference between the present model and most previous models is the focus on present conditions (the stock and its attributes) prior to consideration of additional potential flows from new goods and services (purchase and brand choices).

The exogenous variables include the three sets of "inputs" in the Engel, Kollat and Blackwell (1968) model. All personality variables, such as the need for achievement (Gardner, 1972), are included as predispositions. Socioeconomic and demographic variables are included as constraints, while aggregate supply and demand and prices are included as market factors.

The rationale for using such a global concept as general satisfaction includes the assumption that any purchasing behavior would arise from actual or threatened loss of life satisfaction. An important question for the individual is which deficit is most important and most in need of reduction, taking into account the socioeconomic and other constraints. The particular goods (including brand choice) to be purchased would depend upon the answer to a two part question: (1) what is the greatest deficit and (2) which goods and/or services have the characteristics which will best reduce the deficit (Lancaster, 1971; Maynes, 1973).

The model is extended to the prediction of behavior that might be consequent upon a rise in the total deficits and a reduction in life satisfaction. Being in the higher or lower range of total normative deficits would have implications not only for satisfaction but should produce higher or lower levels of motivation to reduce the normative deficits. Presumably, if satisfaction is reduced sufficiently and motivation to react is high enough, behavior to reduce deficits should appear. That behavior, of course, may be more or less successful in reducing the total normative deficit.

Over time, the degree of success of the consequent behavior should have feedback effects on the other variables in the system. More correctly, successful behavior at  $t_1$  should raise the predisposition to engage in behavior to reduce deficits and improve satisfaction at  $t_2$ . Thus, the effect is not strictly feedback, but a chaining forward in time of the five levels of variables. The behavior might alter the values of the original exogenous variables. A move to a different residence might reduce normative deficits by improving job opportunities and thereby raise the family's income. Higher income would serve to reduce the budget constraint and thereby permit a reduction in one or more other deficits.

A series of successful actions should produce a rise in the individual's (family's) standards by raising their expectations for success in the future (Tallman, 1971). A series of failures might lead to feelings of powerlessness and apathy, and a lowering of personal standards.

#### The Classes of Variables

Five sets of variables are included in the proposed model: the normative deficits, general satisfaction, the exogenous variables, the propensity to act, and consequent behavior.

#### The Normative Deficits

The quality of life may be conceived as a low level of deficits in all consumption areas. That is, a summation of departures of achieved or actual states from (1) personally desired states and (2) culturally or externally prescribed states. Typically both a personal norm and a cultural norm would exist for each attribute of a given consumer good. There may be attributes for which some individuals or even all individuals have no standards. Likewise, there may be areas for which no cultural norm exists. One of the expected products of research to test the model, therefore, would be a set of useful external standards, departures from which would relate to satisfaction, motivation and behavior. Such standards would be useful in developing social policy.

The measurement of specific personal deficits requires attention to both the breadth of the deficit and the base point (or actual state), since given deficits at various levels may have different implications for satisfaction and consequent behavior. For example, a deficit of three bedrooms in a family's housing when the current house has six bedrooms is necessarily different from a three bedroom deficit when the current house has only two bedrooms.

The measurement model for total normative deficits based on personal standards ( $TND_p$ ) is:

$$TND_p = D_{pi} \cdot w_{pi} | base_{pi} + D_{pi+1} \cdot w_{pi+1} | base_{pi+1} \dots \quad (2)$$

where:  $D_{pi}$  = the difference between a personal norm and the actual state for the  $i^{th}$  attribute of some consumer good  
 $base_{pi}$  = the actual state for the  $i^{th}$  attribute  
 $w_{pi}$  = the weight given the  $i^{th}$  attribute by the individual

A basepoint is not needed for measurement of some deficits as culturally prescribed when a common standard is applied to all families. Exceptions would be (1) deficits that have interfamilial contingencies for calculations, as the bedroom deficit variable, which

depends upon family size and composition; (2) time series analysis in which the cultural standards may be changing over time; and (3) cross-cultural research in which the standards differ between cultures.

The measurement model for the cultural total normative deficits ( $TND_c$ ) is:

$$TND_c = D_{ci} \cdot w_{ci} | base_{ci} + D_{ci+1} \cdot w_{ci+1} + D_{ci+2} \cdot w_{ci+2} \dots \quad (3)$$

where:  $D_{ci}$  = the difference between a cultural norm and actual state for the  $i^{th}$  attribute of some consumer good  
 $w_{ci}$  = the weight given the  $i^{th}$  attribute by the culture  
 note: not all cultural deficits include a base point

#### General Satisfaction

General life satisfaction can be viewed as a function of the summed personal deficits, the summed cultural deficits and the exogenous variables:

$$S = f(TND_p, TND_c, X_1, X_2, X_3) \quad (4)$$

It is anticipated that considerable care will be necessary in measuring the concept of satisfaction when the model is tested in light of two factors: (1) the tendency on the part of respondents to prefer to respond in the positive range (satisfied) to any question on satisfaction, and (2) the relatively weak relationships found in many studies between consequent behavior and satisfaction.

#### Exogenous Variables

The exogenous variables are grouped into three classes. The first class may best be termed socioeconomic constraints. Included are such variables as race which, because of discrimination, may increase the probability that members of particular racial and ethnic groups would have high normative deficits. Income, in the form of a budget constraint, may raise the total deficit a family may experience.

The second class of variables includes market factors: supply, demand, and their effect on price and the availability of goods and services. The third class of exogenous variables includes value orientations, world views, and predispositions of various kinds, such as apathy, alienation, the need for achievement, etc. When the model is taken dynamically, the exogenous variables may be influenced by feedback from changes in the level of deficits, satisfaction and behavior.

#### The Propensity to React and Consequent Behavior

The concept, propensity to react, is included to provide a measure of the desires and intentions to engage in behavior that might reduce deficits. Other variables controlled, dissatisfaction should have a predictable relationship with desires and intentions to act to reduce the dissatisfaction, the source of which is a high level of normative deficits. Thus, an expectation or a desire to move to a new neighborhood should result from dissatisfaction with deficits arising from the character of the current neighborhood.

A final assessment of the impact of exogenous variables, normative deficits and satisfaction on family behavior would be the consequent behavior itself.

## Equations and Hypotheses

The set of equations to be estimated are as follows:

$$CB = a + bPB + bCON_1 + bMAR_1 \quad (5)$$

$$PB = a + bSAT + bCON_0 + bMAR_0 + bPRED_0 \quad (6)$$

$$SAT = a + bTND_p + bTND_c + bCON_0 + bMAR_0 + bPRED_0 \quad (7)$$

$$TND_c = a + bCON_0 + bMAR_0 \quad (8)$$

$$TND_p = a + bCON_0 + bMAR_0 + bPRED_0 \quad (9)$$

where:

CB = predicted consequent behavior

PB = predicted propensity to behave

SAT = predicted satisfaction

TND<sub>c</sub> = predicted total cultural deficit

TND<sub>p</sub> = predicted total personal deficit

CON<sub>1</sub> = socioeconomic constraints at t<sub>1</sub>

MAR<sub>1</sub> = market factors at t<sub>1</sub>

SAT = general satisfaction

CON<sub>0</sub> = socioeconomic constraints at t<sub>0</sub>

MAR<sub>0</sub> = market factors at t<sub>0</sub>

PRED<sub>0</sub> = psychological predispositions at t<sub>0</sub>

The a's and b's are constants derived from regression. The beta coefficients for the above system of equations may need to be estimated by a two-stage least squares procedure in order to overcome identification problems.

The relationships shown in the figure are lettered from a to q. Following each letter in parentheses is the hypothesized direction of the relationship.

Equation (5) with the consequent behavior as the dependent variable consists of three direct relationships shown in figure 1 (a,b,c). The key relationship in equation (5) is from the propensity variable which consists of desires and expectations with respect to some behavior that would reduce one or more deficits. The translation of a behavioral propensity into actual behavior is modified by the restraining effects of the concurrent (t<sub>1</sub>) market and socioeconomic constraints. All remaining variables in the model are hypothesized to relate to behavior indirectly through their influence on the propensity variable.

Equation (6) predicting the propensity to act consists of four direct relationships (d,e,f,g). The primary one is general satisfaction. Satisfaction would have a negative relationship to the propensity to behave. Propensity would be negatively affected by previous experience with the market and socioeconomic constraints (t<sub>0</sub>). In addition, various personality factors would affect the propensity to behave either positively or negatively. A high need for achievement, for example, might produce a greater propensity to behave than a low achievement need with identical levels of satisfaction.

Equation (7) predicting satisfaction includes five relationships (h,i,j,k,l). H and i are of key interest in that general satisfaction is seen to be due primarily to the personal and cultural deficits. There is

the possibility that perceived constraints and market factors could attenuate the decline in satisfaction as deficits become higher. Different psychological predispositions may affect satisfaction in both directions. Apathy for example, might reduce the tendency to become dissatisfied.

Equation (8) is a prediction of the total cultural deficit from two of the sets of exogenous variables. The constraints and market factors (t<sub>0</sub>) operate to determine the level of actual or achieved consumption for comparison with cultural norms. The cultural norms are given and for present purposes are assumed not to change. Equation (9) is more complex than (8) because the psychological predispositions are seen to be responsible for the personal standards to the extent that they differ from the cultural norms.

## Discussion

A model such as is presented here could emphasize several levels of behavior and attitude formation. Brand choice, choices among competing classes of goods and attitudes, intentions and expectations, etc. with respect to those behaviors could be studied. Further, there are important decisions having to do with choices of types of behavior that might serve as alternatives to purchasing behavior. Clothing can be home sewn, stolen, or obtained free from charitable organizations.

The model as developed to this stage stops at the point where the family decides whether to engage in consumption behavior in response to their deficit-produced dissatisfaction. Still to be dealt with are such questions as (1) what is needed, food, clothing, housing etc. (obviously, this decision would be based on the deficits that produced the largest amount of dissatisfaction within the constraints of the market and socioeconomic factors), (2) whether to buy pants or dresses and (3) whether to buy them at Sears or Wards. It would appear that the more general decisions would be most influenced by stable cultural norms while brand choice and store choice would depend upon family norms and less stable cultural norms such as fads and fashions.

It is obvious that the development of models that adequately explain consumer behavior requires an approach that is complex in two senses. First, the attributes of the consumption goods that are sought by even very low income families are numerous. It is necessary not only to consider these attributes but also to consider affective, cognitive and other aspects of the attitudes held toward those attributes. Second, the factors that need to be considered in addition to the attributes and attitudes toward them are numerous. A statistical analysis that can be trusted to sort all of that out and produce results that can be managed has not been invented as yet. Nevertheless, partial approaches that represent attempts to comprehend some of the complexity and richness of the socialpsychological and cultural determinants of consumer behavior must be made.

## References

- James R. Bettman, Noel Capon and Richard J. Lutz, "Multiattribute Measurement Models and Multiattribute Attitude Theory: A Test of Construct Validity," Consumer Research, 1(March 1975), 1-15.
- Ann F. Bresler, "Residential Crowding, Fertility Limitation and the Propensity To Move," M. S. thesis, Iowa State University, Ames, Iowa, 1975.

Carol A. Bross, "Normative Housing Deficits and Residential Adaptation," M. S. thesis, Iowa State University, Ames, Iowa, 1975.

James F. Engel, David T. Kollat and Roger D. Blackwell, Consumer Behavior (New York: Holt, Rinehart and Winston, 1968).

Martin Fishbein, "An Investigation of the Relationships Between Beliefs About An Object and the Attitude Toward That Object," Human Relations, 16(August, 1963), 233-240.

David M. Gardner, "An Exploratory Investigation of Achievement Motivation Effects on Consumer Behavior," Proceedings, Third Annual Conference of the Association for Consumer Research, 1972, 20-33.

Peter M. Gladhart, "Family Housing Adjustment and the Theory of Residential Mobility: A Temporal Analysis of Family Residential Histories," Ph. D. dissertation, Cornell University, Ithaca, New York, 1973.

Reuben Hill and Donald A. Hansen, "The Identification of Conceptual Frameworks Utilized in Family Study," Journal of Marriage and Family Living, 22(November, 1960), 299-311.

Reuben Hill and Roy H. Rodgers, "The Developmental Approach," in Harold T. Christensen, Handbook of Marriage and the Family (Chicago: Rand McNally, 1964), 171-211.

John A. Howard, Marketing Management Analysis and Planning, rev. ed. (Homewood, Ill.: Richard D. Irwin, 1963).

David T. Kollat, Roger D. Blackwell and James F. Engel, "The Current Status of Consumer Behavior Research: Developments During the 1968-1972 Period," Proceedings, Third Annual Conference of the Association for Consumer Research, 1972, 576-585.

Kelvin Lancaster, Consumer Demand (New York: Columbia University Press, 1971).

E. Scott Maynes, Decision-making for Consumers, unpublished manuscript (St. Paul: University of Minnesota, 1973).

Earl W. Morris, Sue R. Crull and Mary Winter, "Housing Norms, Housing Satisfaction and the Propensity to Move," Journal of Marriage and the Family, forthcoming.

Earl W. Morris and Peter M. Gladhart, "Mobility, Fertility and Residential Crowding," unpublished paper, Home Economics Research Institute, Iowa State University, Ames, Iowa, 1975.

Earl W. Morris and Mary Winter, "A Theory of Family Housing Adjustment," Journal of Marriage and the Family, 37(February, 1975), 79-88.

Masao Nakanishi and James R. Bettman, "Attitude Models Revisited: An Individual Level Analysis," Consumer Research, 1(December, 1974), 16-21.

Francesco M. Nicosia, Consumer Decision Processes: Marketing and Advertising Implications (Englewood Cliffs, N. J.: Prentice-Hall, 1966)

Richard W. Olshavsky and John O. Summers, "A Study of the Role of Beliefs and Intentions In Consistency Restoration," Consumer Research, 1(June, 1974), 63-70.

Talcott Parsons and E. A. Shils (eds.), Toward A General Theory of Action (Cambridge: Harvard

University Press, 1954).

Roy H. Rodgers, Improvements in the Construction and Analysis of Family Life Cycle Categories (Kalamazoo: Western Michigan University, 1962).

Roy H. Rodgers, "Toward A Theory of Family Development," Journal of Marriage and the Family, 26(August, 1964), 262-270.

Milton J. Rosenberg, "Cognitive Structure and Attitudinal Affect," Journal of Abnormal and Social Psychology, 53(November, 1956), 367-372.

Irving Tallman, "Family Problem Solving and Social Problems," in Joan Aldous, Thomas Condon, Reuben Hill, Murray Straus and Irving Tallman's Family Problem Solving (Hinsdale, Ill.: Dryden Press, 1971), 324-350.

Gabriel Tarde, The Laws of Imitation (New York: Henry Holt, 1903).

Scott Ward, "Consumer Socialization," Consumer Research, 1(September, 1974), 1-14.

David T. Wilson, H. Lee Mathews and James W. Harvey, "An Empirical Test of the Fishbein Behavioral Intention Model," Consumer Research, 1(March, 1975), 39-48.

COMMENTS ON A STOCHASTIC INVENTORY MODEL  
AND A NORMATIVE DEFICIT MODEL

Richard F. Yalch, University of Washington

The two papers by Beutler and Morris, Winter, and Beutler present novel perspectives of consumer behavior. In commenting on their contributions, I will attempt to identify the strengths and weaknesses to assist those interested in applying the models or further pursuing these lines of inquiry.

#### A Stochastic Inventory Model

Beutler addresses the largely ignored issue of how consumers manage their financial resources in anticipation of uncertain incomes and expenditures. He applies a model previously developed to explain a business firm's asset management and demonstrates the potential utility of many financial and economic models. Three aspects of his model will be discussed, the appropriateness of the assumptions, the validation of the model, and the possibility of modifying the model to bring it more in line with extant consumer behavior research findings.

#### Appropriateness of the Assumptions

Although Beutler claims that he is able to derive the cost minimization decisions with only three basic assumptions, there are in fact quite a few more. It is important that these be acknowledged to forewarn anyone seeking to apply the model and to guide those interested in extending it. The first assumption, each household has an asset manager who acts to minimize the expected cost of managing the household's assets by following certain policy rules, actually assumes several behaviors. The first is that the household members manage their finances in concert or with one individual in control. Research suggests that purchases are typically divided according to interest and expertise (cf. Ferber & Lee, 1974), such that there may not be much coordination. Another, that the asset manager follows an (h-z) policy, requires that the cash balance falls to zero before there will be a transfer from short-term assets. Given the propensity for impulse buying, it would seem more likely that a positive balance would be desired at all times. Also, the assumption that the consumer is indifferent between holding currency and demand deposits is tenuous and neglects a decision of great interest to bank managers and economists.

Beutler's second major assumption is that asset inventories are stochastic, which is later modified to require equally likely increases or decreases in the cash balance at any point in time. This limits the model to families with nonpersistent income changes and is at odds with the fact that most families receive their income at known predetermined dates, i.e., end of the month.

On the basis of the first two explicit and many implicit assumptions, Beutler develops a model which differs from Miller and Orr's (1966) in only two ways. One is that a constant is added to the short-term asset balance,  $Z_0$ , to reflect a desire to maintain a positive balance but no reason is given for this desire. The other modification is that an independent variance is added to the short-term asset balance as the means to treat unexpected gains and savings for large purchases.

It is not clear why windfall gains would not be converted to long term assets.

The third major assumption is that the asset manager will act to minimize expected costs, which was also stated under assumption one. Beutler assumes that individuals are aware of the costs of holding different assets, but research suggests that this is rarely true (White, 1975). If consumers are not cognizant of the costs, it hardly seems likely that they will be able to minimize them. Beutler also assumes that consumers will consider the holding of real assets as costing the same as cash. A better approach might be to consider the investment motive for durable good purchases and treat it the same as long-term assets.

#### Model Validation

Beutler does not empirically validate his model but merely demonstrates that certain changes in the parameters would lead to the proper predicted behaviors. This is a reasonable first step in evaluating a simulation model, but is only a necessary and not sufficient condition for acceptance. Furthermore, since the model was not derived from a careful consideration of prior consumer behavior research findings, an empirical test would seem particularly desirable.

#### Extensions

The following modifications are proposed to bring the model more in line with our knowledge of consumer behavior. The effect these suggestions might have on the mathematical simplicity, apparently a major consideration in the present formulation, is not known but there is a clear tradeoff between developing a workable model and modelling the world as it really behaves.

1. Allow consumers to maintain a positive cash balance instead of requiring it to drop to zero before transferring short-term asset funds.
2. Separate the holding of currency from demand deposits by assigning different costs to each.
3. Consider what effect nonstationary incomes will have on the optimum levels.
4. Allow the minimum asset balance,  $Z_0$ , to fluctuate with environmental changes.
5. Establish a more realistic cost to holding real assets.

#### A Normative Deficit Model

Morris, Winter, and Beutler criticize extant consumer behavior research because of the "narrow social psychological focus," failure to specify operationalizations, concern with selective demand, exclusion of cultural influences, and utilization of product attributes rather than consumer needs as the behavior determinants. These criticisms serve as the motivation for offering yet another model purported to explain and predict consumer behavior. In evaluating this model, it would be beneficial to consider the validity of their criticisms and their model's ability to overcome them.

#### Narrow Social Psychological Focus

It is true that social psychology has had the greatest impact on consumer behavior research, which is understandable given the problem setting, individuals making decisions on the basis of information and prior experiences. In fact, since the authors maintain the individual as the unit of analysis, it is not clear that their approach is significantly different. Also, the authors seem to misrepresent the limits of this approach. For example, although attitudes are frequently considered the most important explanatory variable, they are rarely the only variable considered. The recent interest in extended prediction models (Ryan & Bonfield, 1975) and situational factors (Belk, 1974) reflect this trend. Furthermore, the authors' proposal that consumers select products on the basis of their relative abilities to satisfy needs would appear to be very similar to the inclusion of an ideal product in a multiattribute or multidimensional scaling analysis of consumer attitudes.

#### Operationalization

Morris *et al.* are critical of the integrative consumer behavior models because the authors have neglected to specify exactly how the variables and relationships should be operationalized. Unfortunately, they are equally guilty of this and are even vague as to what is included in the model let alone indicating how it might be tested. For example, the exogenous variable, market factors, is defined as, "supply, demand, and their effect on price and the availability of goods and services." This could be interpreted as stating that consumer behavior is determined by the marketing activities of the firms in an industry, hardly a noteworthy contribution. Moreover, the analytical method for estimating the relative influence of each variable is equally unclear. With lagged terms, two-way causation (consequent behavior is assumed to influence predispositions), and direct and indirect influences, this would not seem to be an easy task.

#### Product Choice Decision

In considering the decision to purchase a product instead of the choice of one brand over another, the authors point out that researchers have neglected an important aspect of consumer behavior. Apparently, many feel that the brand choice decision is of more importance since firms benefit directly by stimulating selective demand and only partially from primary demand changes. Recent concerns over inherent shortages and the consumption of undesirable products justify paying more attention to primary demand.

Unfortunately, the authors may be overstating their model's ability to explain primary demand decisions because of prior work on house buying. The assumption that the purchase of nondurable goods also comes about after a consideration of which satisfies the greatest need may not be valid. Consumers may be torn between purchasing a new house or car, but it is doubtful whether they consider toothpastes, haircuts, and cereals in the same manner. The authors might be wise to limit their model to the choice between consumer durable goods.

#### Cultural Norms

The authors argue that product choices are based on the relevant cultural norms as well as one's own desires. That is, one might desire a game room because others think it is important to have. This appears to be similar to Fishbein's (1967) proposition that intentions are based partially on social normative beliefs. Questions arise as to whether these are best measured by asking the individual for his perception of the cultural norms or to try to assess them in some other

manner. The former has the advantage in that you can be certain that the individual is at least aware of the norm though responses may be biased by personal norms. The latter eliminates the bias but creates a problem in that many of the measured values may not be known by the subject. If so, it should not be considered as having an influence.

#### Consumer Needs

Recent consumer behavior research has been dominated by efforts to explain attitudes and behaviors on the basis of product attributes. Morris *et al.* propose that the true influence is the need that might be satisfied by the purchase of the product. However, in discussing their model they tend to focus on attributes rather than needs. For example, deficits are described in terms of the number of bedrooms needed not the sleeping needs that must be satisfied. The latter would allow the purchase of a sofa bed as a solution but the former would not. Thus, they have not followed through on their own recommendation.

#### Summary

Beutler has contributed to consumer behavior research by illustrating the close correspondence to business decision making, investigating savings rather than purchases, using finance as the theoretical framework, and developing a normative model. In evaluating this noteworthy effort, one should be aware of the many simplifying assumptions. Morris *et al.* have proposed a model that is consistent with the marketing concept in that need satisfaction is viewed as the motivation for product choices. The potential contribution should be greatest for consumer durable good purchases. Both models await empirical validation.

#### References

- R. W. Belk, "An Exploratory Assessment of Situational Effects in Buyer Behavior," Journal of Marketing Research, 11 (May 1974), 156-163.
- R. Ferber and L. C. Lee, "Husband-Wife Influence in Family Purchasing Behavior," Journal of Consumer Research, 1 (June 1974), 43-50.
- M. Fishbein, "Attitude and the Prediction of Behavior," in M. Fishbein, Ed., Readings in Attitude Theory and Measurement. (New York: Wiley, 1967), 477-492.
- M. H. Miller and D. Orr, "A Model for the Demand for Money by Firms," Quarterly Journal of Economics, 80 (1966), 413-435.
- M. J. Ryan and E. H. Bonfield, "The Fishbein Extended Model and Consumer Behavior," Journal of Consumer Research, 2 (September 1975), 118-136.
- K. J. White, "Consumer Choice and Use of Bank Credit Cards," Journal of Consumer Research, 2 (June 1975), 10-18.

EFFECTS OF EXPECTATION CREATION AND DISCONFIRMATION  
ON BELIEF ELEMENTS OF COGNITIVE STRUCTURE<sup>1</sup>

Jerry C. Olson, The Pennsylvania State University  
Philip Dover (student), The Pennsylvania State University

Abstract

Previous research on consumer expectations is briefly reviewed. A major problem in this work is the lack of a precise and theoretically rich conceptual definition of expectations, since most other problems stem from the resultant ambiguity regarding the expectancy construct. It is suggested that expectations are best considered as individual belief elements in a cognitive structure. In a longitudinal experiment, expectations about an attribute of a new brand of coffee were created by exposure to three ad-like messages and then were disconfirmed through actual product trial. The effects of these manipulations on belief-expectancies and other elements of cognitive structure are examined, as are various theoretical perspectives for future research.

Much has been written about the importance of consumer expectations, particularly their impact on purchasing behavior and the national economy (cf. Katona, 1974; Kelley and Scheewe, 1975) and their effects on consumer dissatisfaction (cf. Anderson and Hair, 1972). However, despite the seeming relevance of expectancies to several major issues, only a handful of empirical studies in the marketing literature deal explicitly with consumer expectations (Anderson, 1973; Cardozo, 1965; Cohen and Goldberg, 1970; Olshavsky and Miller, 1972; Swan, 1972; Woodside, 1972). Most of this research has concentrated on the effects of a product usage experience that disconfirms pre-trial expectations.

In a critical review of this research, Olson, Toy and Dover (1975) identified several theoretical and methodological problems which may be inhibiting consumer behaviorists from examining important expectancy phenomena. The present paper summarizes these problem areas and describes an experimental study which focuses on the cognitive processes underlying expectancy formation and disconfirmation. This research was designed to eliminate many of the problems found in earlier studies and to explore theoretical bases for future work.

Problems in Previous Research

Manipulation of Expectations

In each of the marketing expectancy studies, product expectations were supposedly created by providing consumers with product attribute and/or performance information prior to actual trial. The usual manipulation involved a written communication about specific characteristics of the product. For instance, Olshavsky and Miller (1972) presented consumers with a list of features about audio tape decks, including price, brand name, and several technical/performance attributes. Similarly broad informational manipulations of consumer expectations were used by Cardozo (1965),

Woodside (1972), and Anderson (1973).

These "shotgun approaches" to expectancy creation indicate a conceptual vagueness regarding the expectation construct. In none of the studies did the author(s) specify precisely what type of expectation was to be created by the manipulation used. Moreover, logical or theoretical rationales were not offered regarding why the manipulation should be effective. Rather, it seemed to be implicitly assumed that exposure to information about product attributes would, by some unspecified process, create an expectancy regarding the product. Other evidence for the conceptual uncertainty regarding the expectancy construct is the almost total lack of measures to determine whether or not the informational manipulations actually created expectations (only Anderson, 1973, used manipulation checks). Of course, empirical manipulation checks are difficult to design without a precise idea of what is being manipulated.

Three other problems involving the expectancy-creating manipulations should be briefly noted. First, there was a general failure in the past research to identify and use alternative external sources of consumer expectations. In addition to general written information about product features, product expectations may be created by exposure to advertisements and word-of-mouth, by observation of the product, and through actual product usage experiences. Second, although expectancies are probably formed most often by a series of exposures to such sources over time, none of the expectancy research incorporated this longitudinal dimension. Third, the expectancy manipulation and its position within the overall experimental procedures typically were somewhat artificial and transparent. For instance, in all of the studies except Cohen and Goldberg (1970), consumer subjects were given the informational expectancy manipulation immediately prior to the product usage experience which in turn was immediately followed by the post-trial questionnaire. Such short time periods between experimental events are somewhat artificial and also tend to increase the transparency of the experimental purpose, perhaps creating demand characteristic artifacts. None of the studies reported use of debriefings or other procedures to determine whether or not some respondents had discovered the experimental purpose.

Theoretical Perspective

Much of the expectancy research may be criticized for a lack of attention to alternative theoretical bases for disconfirmation phenomena. Beyond the typical broad references to dissonance theory, only Anderson (1973) examined various alternative theoretical explanations for his results.

Conceptual Definition of Expectation

A major problem in the consumer expectancy literature, and a primary cause of the problems discussed above, is the lack of a conceptually precise definition of an "expectation." Partly because of the resultant ambiguity regarding the expectancy construct, investigators have used various manipulations and theoretical perspectives in their research. However, broadly speaking,

<sup>1</sup>This research was funded by a grant from the Pennsylvania State University to the senior author and Andrew Mitchell.

in all of the previous studies expectancies seem to have been inferred from consumers' general, overall evaluative judgments of the product. Examples of this "conceptual approach" include Cardozo's overall rating of product value, Cohen and Goldberg's overall product preference rating, Olshavsky and Miller's rating of overall product performance, and Anderson's overall product rating and aggregated index based on separate ratings of 15 product attributes. Thus, although explicit definitions of expectancies were not found in the previous research literature, expectations seem to have been implicitly conceptualized as global product evaluations. Most theorists agree that such evaluative judgments are attitudes toward the product (cf. Fishbein and Ajzen, 1975). Thus, the past research has implicitly considered consumer expectations as general, overall product attitudes.

We believe this conceptual perspective to be somewhat vague, potentially misleading, and therefore dysfunctional. By considering expectancies as global product attitudes, researchers tend to ignore the cognitive processes underlying attitude formation and attitude change due to a disconfirmation experience. In order to focus research on these cognitive attitude processes, the expectation construct must be given a precise meaning that will unambiguously identify its cognitive nature and position it theoretically vis-a-vis other cognitive elements. To that end, we recommend that consumer product expectations be conceptualized as individual belief elements in a consumer's cognitive structure regarding the product.

As a scientific construct, belief has a precise and generally accepted definition; that is, one's subjective judgment regarding the likelihood of a relationship between two concepts (cf. Fishbein, 1967). Thus, an expectation is the perceived likelihood that a product possesses a certain characteristic or attribute, or will lead to a particular event or outcome. Essentially, beliefs and expectancies are the same cognitive construct, although some may wish to use the term expectation only when referring to a belief judgment regarding a "future event or state of affairs." We should note that this conceptual approach is not new; for example, Tolman (1932) considered expectancies to be "cognitions" or "subjective probabilities that one event is associated with (or follows from) some other event (Fishbein and Ajzen, 1975, p. 28)."

By considering expectancies as beliefs, research on consumer expectations is placed squarely into the theoretically-rich literature dealing with cognitive structure (cf. Bettman, Capon and Lutz, 1975; Bither and Miller, 1968) and attitude formation and change (cf. Lutz, 1975; Olson and Mitchell, 1974). Such a conceptual perspective is useful in several ways. First, based upon the fairly extensive body of research which has examined belief formation processes, manipulations of consumer belief-expectations may be more easily operationalized. Second, a cognitive perspective to expectancy research forces attention on the elements of cognitive structure (e.g., belief-expectations, the evaluative aspect of belief-expectancies, overall evaluation or attitudes toward the product, behavioral intention, etc.), and on their interrelationships. Third, based upon the massive literature dealing with the relationships between beliefs, attitudes, behavioral intention, and actual behavior, consumer expectations may be theoretically linked with subsequent overt behaviors such as purchase. Fourth, by considering expectations as cognitive beliefs, various cognitive theories or learning theory may be used to explain the acquisition of overall product attitude (Olson and Mitchell, 1974). Then, if desired, this attitude may be considered as a summary or global expectation about the product.

## Purpose of Present Research

The present study represents an initial examination of consumer expectations as attribute-specific product beliefs and of the disconfirmation of expectations from the more theoretically justifiable perspective of cognitive structure modification. This conceptual approach leads to an explicit consideration of the processes by which belief-expectations are created and of the effects that such a cognitive structure has on product judgments following a disconfirming usage experience. Clearly, this conceptual perspective requires a longitudinal research design in which each subject's cognitive structure is monitored over time as a function of various expectancy-creating and expectancy-disconfirming informational inputs. The present longitudinal, individual-level approach contrasts with the earlier research, all of which has used between-groups, aggregate data analyses. In summary, the present research is an attempt to operationalize the suggestions presented above and, particularly, to illustrate the advantages of a more cognitively oriented approach to research concerning the acquisition and disconfirmation of consumer expectations. Furthermore, although competing theoretical explanations for disconfirmation effects are not explicitly tested in this study, a second purpose is to examine relationships between selected constructs from which useful theoretical bases for future research may be inductively derived.

## Methods

### Research Design

The primary purpose of this research was to examine the process of expectancy creation and the subsequent effect of these beliefs on product evaluations following an expectancy-disconfirming usage experience. Therefore, the major manipulation was four "trials" in which consumer subjects were exposed to product information that could be acquired and integrated into their cognitive structures (see the research design presented in Table 1). In trials 1, 2, and 3, subjects in the experimental group were presented with three different, ad-like communications, each of which stressed that the product

TABLE 1  
EXPERIMENTAL DESIGN

	Trials (Information Exposure)			
	1	2	3	4
Treatment Groups	"Ad"A	"Ad"B	"Ad"C	Prod. Use
Experimental Group (n=20)	X-0	X-0	X-0	X-0
Control Group (n=18)				X-0

Note--"X" denotes exposure to information; "0" denotes observation or measurement, as in Campbell and Stanley (1963).

possessed a specific product characteristic and which should create belief-expectancies to that effect. Trial 4 consisted of an actual product usage experience which



was purposefully devised so as to disconfirm the manipulated expectancy beliefs. A separate control group received none of the three expectancy-creating messages, but did try the product in the same usage experience setting. Following each informational exposure (i.e., each trial), extensive measures of cognitive structure were taken.

Product

Several authors have suggested that some degree of ego-involvement with a product may be necessary in order that disconfirmation effects are clearly evidenced (Cohen and Goldberg, 1970; Anderson, 1973). With this in mind, ground coffee was chosen as a product thought to be relatively high in ego-involvement for the subjects used in this research. Additional considerations also influenced the choice of ground coffee. Coffee has a relatively limited cognitive structure, in terms of the number of salient belief attributes, thus simplifying the measurement requirements. Moreover, the authors had access to a new brand of coffee (of below-average quality) which was not available in the local area and, therefore, was unknown to the subjects. In order to increase the mundane realism of the experiment, the actual brand name of this product was used throughout the study. However, for purposes of this paper, the product is referred to as "Coffee."

Subjects

The subjects for this study were 38 adult women recruited from several small social groups which met regularly on a bi-weekly basis. Most of these women were married, with children in their family, and all were regular consumers of ground coffee. The 20 women in the experimental group were promised \$5.00 and the 18 control group women were promised \$2.50 for their respective participation. Consumers in both groups were paid upon completion of the study.

Dependent Variables

Although numerous methods for measuring belief elements of cognitive structure have been used in marketing research (e.g., Fishbein and Raven, 1962; Sheth, 1974; Wilkie and Pessemier, 1973), certain dissatisfactions with these procedures exist (cf. Cohen, Fishbein and Ahtola, 1972). Recently, a seemingly improved model and measurement procedure was proposed by Ahtola (1975). This "vector model" maintains the theoretical properties of the belief construct while allowing a more realistic measurement of a consumer's belief structure regarding a product. In order to implement this procedure, detailed knowledge about consumers' cognitive structures is required. The major belief-expectancy dimensions (or salient product attributes) in a cognitive structure, as well as the specific perceptual categories or levels along each dimension, must be determined (cf. Mazis, Ahtola, Klippel, 1975). Once these cognitive aspects are identified, product beliefs are measured not for a specific dimension (e.g., sweetness), but for each discriminable level or category along that dimension (e.g., slightly sweet, very sweet, etc.).

In order to identify these aspects of cognitive structure for ground coffee, extensive pilot interviews were conducted with 15 housewife subjects using a modified repertory grid procedure (for details of this method see Wilson and Dover, 1975). Five attributes of ground coffee and their respective discriminable levels were identified as the most frequently appearing elements of cognitive structure (see in Table 2).

TABLE 2  
BELIEF-EXPECTANCY ELEMENTS OF COGNITIVE STRUCTURE  
FOR GROUND COFFEE

Belief Dimension	Discriminable Belief Levels of Each Attribute Dimension				
Bitterness	not at all bitter	slightly bitter	fairly bitter	very bitter	
Amount of Caffeine	none at all	less than most other brands	about the same	more than other brands	
Relative Expensiveness	much more	a little more	about the same as other brands	a little less	much less
Strength of Flavor	very weak	fairly weak	neither weak nor strong	fairly strong	very strong
Consistency	very consistent	fairly consistent	fairly inconsistent	very inconsistent	

These cognitive dimensions were represented in a questionnaire using the same procedure as Ahtola (1975). Briefly, subjects were to assign 10 points among the 4 or 5 levels of each belief-expectancy dimension, such that the number of points granted to each attribute level represented the strength of belief about the likelihood that the product possessed that particular level of the attribute dimension. Other cognitive elements measured in the questionnaire included confidence in each belief-expectancy rating, the evaluative aspect of each attribute level along each dimension, the overall evaluation of (or the attitude toward) the product, and the behavioral (purchase) intention toward the product.

Expectancy Manipulations

The specific belief expectancy dimension chosen for manipulation was coffee bitterness. In the repertory grid pretesting, bitterness generally seemed to be the most important characteristic of ground coffee, with "not bitter," of course, the most desirable level of bitterness. Further pretesting suggested that coffee drinkers could best distinguish four levels along the bitterness dimension--not at all, somewhat, fairly, and very bitter (see Table 2).

Based upon this knowledge, three different messages were created using an "every-day" informal writing style to simulate advertising copy. Each of these messages emphasized that "Coffee" had no bitterness. Because other product attributes were not explicitly mentioned in the messages, only expectancies regarding coffee bitterness were directly manipulated. Each message was typed on official-appearing stationery made up with the company logo as a letterhead. Communication A simulated a typical consumer print ad and described the product in general non-technical terms emphasizing its lack of bitterness. Communications B and C were slightly more technical discussions of "Coffee's" lack of bitterness. The former summarized the results of a consumer survey and the latter described opinions of expert taster-testers.

## Experimental Procedures

Subjects were first contacted at their respective group meetings and were asked to participate in a study of the effectiveness of marketing procedures used to introduce a new product, in this case a new brand of ground coffee. Interest was high and virtually all those contacted agreed to participate.

Subjects in the experimental group were told that manufacturers often send out information about their products in order to inform consumers of the product and its particular characteristics. Further, they were told that the researchers had obtained several of these communications for a new brand of coffee not yet available in the local area. These subjects were told that the present study was being conducted in order to obtain their reactions to these communications and to the coffee itself. The control group subjects were merely told that a new brand of coffee had been developed, but was not yet available in the local area. They were told that the present study was to obtain their tasting reactions to a sample of the coffee at one of their future meetings.

The study began about two weeks after the initial contact. At that time, the first of the communications was personally delivered to each experimental group respondent, along with the questionnaire containing the dependent variables. Subjects were told to read the communication at their leisure and then immediately to complete the questionnaire and mail it back to the researchers. Four days later, the second communication and an identical questionnaire were delivered to the subjects, followed in four days by the third communication and questionnaire. No problems occurred regarding prompt return of the questionnaires.

Then, approximately four days later, the experimental group met for one of its regular meetings. During the meeting, consumer subjects were taken singly to another room for the tasting of the "Coffee." In order to create an expectancy-disconfirming trial experience with the product, a large pot of coffee was made up using 50% more ground coffee than called for by the package directions. In earlier pilot tests, this composition had been judged to have substantial bitterness.

For the tasting, subjects were seated at a table and given one-half cup of black "Coffee" (i.e., no cream or sugar was allowed). Subjects were told that taste sensitivity was enhanced when cream and/or sugar were omitted. In this way a consistent product trial experience was maintained for all subjects. Subjects were allowed to sip and sample the taste and aroma of the coffee for approximately two minutes and then responded to a questionnaire containing the same cognitive structure measures taken in the previous three trials as well as several other measures not relevant to this paper. Identical tasting procedures and measuring instruments were used for the control group subjects when they met for one of their regular meetings.

All subjects were debriefed individually in their homes approximately two weeks later.

## Results and Discussion

### Evaluation of Demand Characteristic Effects

Longitudinal designs involving multiple exposures and measurements tend to increase the potential for demand characteristic artifacts (Olson, 1974; Sawyer, 1975). In order to roughly determine the biasing potential of such artifacts in the present study, subjects were asked, upon completion of the research, for their opin-

ion of its purpose. All repeated essentially the original explanation given them. No subjects mentioned that the study might deal with the effects of a disconfirmation experience or with the development of product beliefs and their impact on later product judgments.

### Manipulation Checks--Expectations

The first three measurements of belief strength for the "not at all bitter" expectancy may be considered as manipulation checks for the three expectancy-creating communications. On the 10-point scale of belief strength, the mean ratings for the belief expectancy that the "coffee" was "not at all bitter" were 7.45, 6.65, and 6.70, over the first three trials. Although the mean expectancy strength decreased slightly after the first communication, the trend was not significant ( $F=1.28$ ,  $df=2/57$ ,  $p>.10$ ). Thus, the communications were successful in creating a rather strong belief-expectation that the "Coffee" was not bitter.

### Manipulation Checks--Disconfirmation

In fact, however, the "Coffee" was rather bitter as evidenced by the control group's post trial ratings. As can be seen by looking ahead to Figure 2, the control subjects, who had no specific expectancy manipulations prior to tasting, judged the "Coffee" to be more bitter, at all four bitterness belief levels, than did the experimental group. Although only the not-at-all-bitter belief-expectancy yielded a marginally significant difference between the experimental and control groups ( $t=1.68$ ,  $p=.10$ ), the overall pattern of ratings provides indirect evidence that the coffee tasting experience did disconfirm the experimental group subjects' "unrealistic" expectations.

### Effects on Belief Elements of Cognitive Structure

The variety of measures taken over the four experimental trials resulted in a large amount of data regarding several cognitive elements and their interrelationships. However, this paper deals only with those results related to belief-expectancies about "Coffee" bitterness and certain other closely-related belief elements in consumer's cognitive structures.

In order to clearly interpret the effects of the messages and disconfirmation experience on the bitterness belief-expectancies, one must consider the evaluative responses to each bitterness belief level. Table 3 presents the mean evaluation ratings of each level of bitterness (for the generic product, ground coffee) taken at each of the four experimental trials and for the control group (post-trial). Three results in these data deserve mention. First, for the experimental group, the evaluative ratings are quite stable over the four trials. None of the four one-way ANOVA's yielded significant differences. Second, none of the control group's evaluations of the bitterness attribute are significantly different from the grand mean of the experimental group's evaluations (all  $p$ 's  $>.10$ ). Third, there are wide differences in evaluative responses between the four belief levels for coffee bitterness. Based upon the means presented in Table 3, the evaluative aspects ( $a_i$ ) of these four bitterness belief-expectancies can be qualitatively labeled as: not at all bitter--very good, slightly bitter--neutral, fairly bitter--fairly bad, very bitter--extremely bad. These distinct differences in evaluation provide further support for treating belief components of cognitive structure in terms of specific categories, levels, or amounts of each attribute dimension.

Figure 1 presents the mean ratings of belief-expectancy strength for the four bitterness levels as a function

TABLE 3

MEAN EVALUATION LEVELS FOR FOUR BELIEF-EXPECTANCIES REGARDING "COFFEE" BITTERNESS

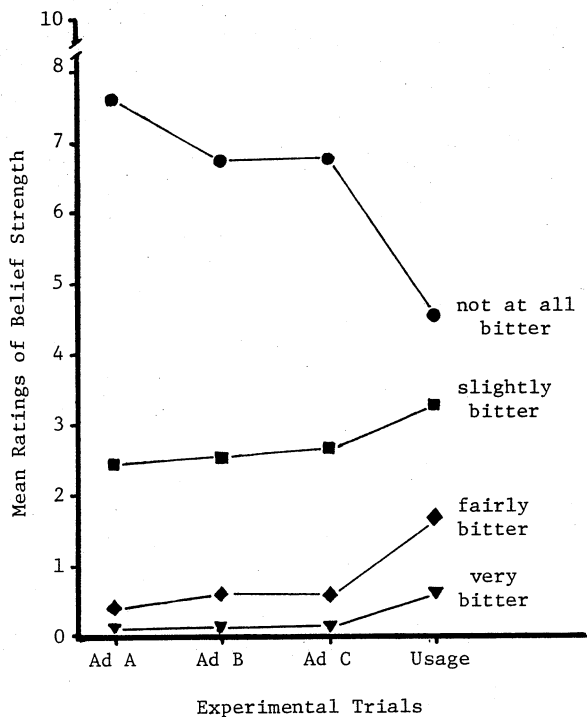
Trial	Rating taken following...	Bitterness Belief-Expectancies			
		Not at all bitter	Slightly bitter	Fairly bitter	Very bitter
1.	"Ad" A	2.10	0.10	-1.45	-2.95
2.	"Ad" B	1.90	0.20	-1.40	-2.80
3.	"Ad" C	2.45	0.10	-1.70	-2.90
4.	Trial	2.55	0.40	-1.30	-2.70
Control	Trial	2.26	0.00	-1.00	-2.70

Note--Evaluation ratings were made on 7-point scales (-3 to +3) labeled "bad" and "good".

of the expectancy-creating messages and the disconfirming trial experience. For each bitterness level the changes in expectancy strength over the four trials were analyzed by a one-way, within-subjects ANOVA.

FIGURE 1

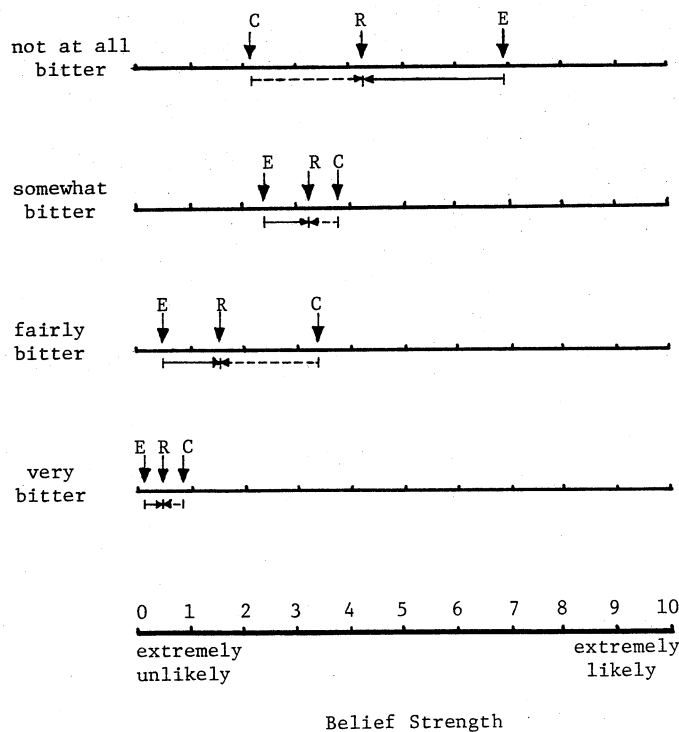
CHANGES IN BELIEF STRENGTH FOR BELIEF-EXPECTANCIES REGARDING COFFEE BITTERNESS



These analyses indicated significant changes in belief-expectations for the "not-at-all-bitter" ( $F=3.29, df=3/57, p<.03$ ) and the "fairly bitter" ( $F=2.61, df=3/57, p<.06$ ) belief levels, but not for the other two bitterness beliefs. Using the Bonferroni method for pairwise comparisons, no significant differences were found between the three pre-trial expectations for the not-at-all-bitter and fairly-bitter belief levels. Therefore, these pre-trial expectancies were averaged and the single belief-expectation was compared with the post-trial belief strength. For both the not-at-all-bitter and fairly-bitter attribute levels, the disconfirming trial experience caused a significant change in belief-expectations (both  $p's<.03$ ). In an evaluative sense, both beliefs changed in more negative direction (i.e., each  $B_{i1}$  combination became more negative, or less positive). In Figure 1 note also that the nonsignificant changes in post-trial belief-expectations for the other two bitterness levels are consistent with the above-mentioned results. Thus, the entire bitterness belief-expectancy vector (i.e.,  $\Sigma B_{i1}$ ) became more negative following the disconfirmation experience. In summary, the overall pattern of belief-expectancy changes is consistent (in an evaluative sense) for all four bitterness levels, with the greatest change in each occurring as a function of the disconfirming usage

FIGURE 2

COGNITIVE CHANGES IN BELIEF-EXPECTANCIES AS A FUNCTION OF A DISCONFIRMING TRIAL EXPERIENCE



Note: E = mean pretrial expectancies--experimental group  
 R = mean post-trial rating--experimental group  
 C = mean post-trial rating--control group  
 - - - = magnitude of assimilation effect  
 — = magnitude of belief change

experience.<sup>2</sup>

Thus, as Cohen and Goldberg (1970) found and as might be intuitively expected, product usage experience appears to have a strong impact on belief-expectancy elements of cognitive structure. These data suggest that the information consumers derive from product trial (or longer-term usage experience) may have a greater effect on the acquisition and/or change of cognitive structure elements than does information from sources such as advertising or word-of-mouth. Of course, variables such as product-specific and general self-confidence could be expected to mediate the cognitive impact of such experientially-derived information.

The changes in belief-expectancies become more clearly interpretable when compared with the belief ratings obtained from the control group, which received no pre-trial manipulation of expectancies regarding the "Coffee." In Figure 2, the consistent effect of the disconfirming trial experience on post-trial belief strength "moved away from" the pretrial expectancy and "towards," but did not achieve, the "objective" post-trial belief strength of the control group. It should be noted that, although the latter result is true in terms of absolute magnitudes of belief strength, only the not-at-all-bitter belief-expectancy evidenced a significant difference in post-trial ratings ( $p=.10$ ) between the experimental and control groups. The direction, if not the magnitude, of each post-trial belief-expectancy change constitutes a perfect example of an assimilation effect (Sherif and Hovland, 1961). In disconfirmation phenomena, assimilation occurs when a stimulus is judged closer to one's own position (pre-trial belief-expectancy) than it objectively is (determined by the control group's post-trial belief strength rating). These findings of consistent assimilation effects are in agreement with the results of virtually all the earlier disconfirmation research in marketing.

However, before concluding that assimilation-contrast is the theoretical basis for explaining disconfirmation effects, it must be noted that at least one alternative theory--namely, cognitive dissonance--also yields predictions that are consistent with the pattern of results reported above and in earlier research. Interestingly, however, the two theories do differ in terms of specific cognitive phenomena of primary interest in a disconfirmation study. For instance, given data as presented in Figure 2, assimilation-contrast theory explains the post-trial belief "R" as a movement from "C" (the "objective" post-trial belief of the control group) caused by the pretrial belief-expectancy "E", whereas dissonance theory focuses on the "E-to-R" belief changes and requires no control group "C".

Because dissonance and assimilation-contrast theories generally predict the same pattern of results, it is quite difficult to devise a clear, distinguishing test between the two. Only when expectations are very strongly disconfirmed do the two theories generate differing predictions. For disconfirmations of great magnitude, assimilation-contrast theory predicts a contrast effect--i.e., a post-trial belief rating "further away from" one's expectation than an objective belief judgment. In comparison, dissonance theory would continue to predict a post-trial belief "closer" to one's expectation than an objective judgment. It is interesting to note that no marketing-related study has found

<sup>2</sup>Although space does not allow a thorough discussion, it is interesting to note that all the other cognitive structure elements measured (e.g., attitude--A<sub>0</sub>, behavioral intent--BI, and the belief vectors themselves--EB<sub>1,1</sub>) evidenced this same pattern of change.

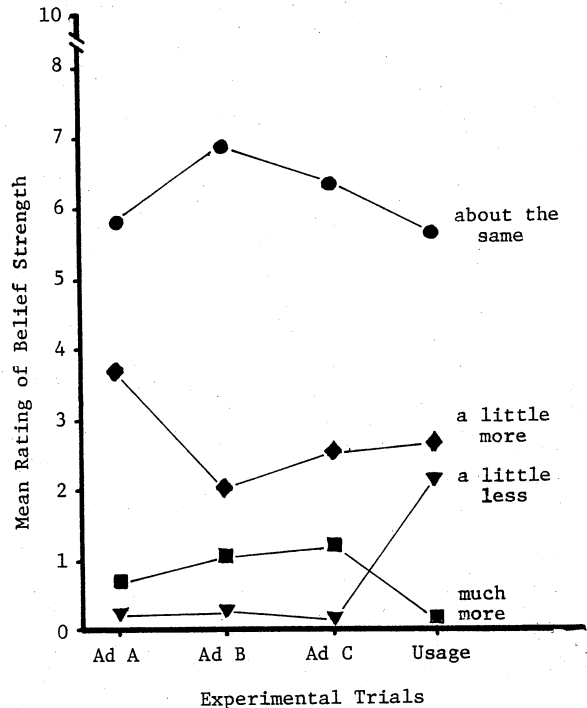
evidence for a contrast effect, although Anderson (1973) incorrectly identified a contrast effect in his extreme disconfirmation condition (see Olson, Toy and Dover, 1975).

Reasoning inductively, the present and earlier results suggest that usage experience with most consumer products may not create the extreme disconfirmations necessary to produce a contrast effect. Alternatively, it may be that post-trial cognitive changes following a disconfirming experience are actually caused by dissonance reduction rather than assimilation processes. Mediating or moderating variables which may be useful in determining whether or not contrast effects can be obtained in consumer disconfirmation situations include: ego-involvement with the product, effort expended and involvement in the choice-usage situation, and the general ambiguity of the product judgments.

A rather different conceptual perspective for viewing disconfirmation phenomena might be derived from the various information processing models of the information integration process. The theoretical alternatives include information averaging models (cf. Anderson, 1970; Bettman, Capon and Lutz, 1975), congruity theory (Osgood and Tannenbaum, 1955), attitude models (cf. Fishbein, 1963; Rosenberg, 1956), and balance theory (Heider, 1946; Newcomb, 1953), among others. Most of the extant research involving these models has examined the outcomes of integrating information about different stimulus attributes (e.g., price and style) or different cognitive elements (e.g., affective and cognitive). However, it may be possible to modify or extend these theoretical approaches in order to handle situations involving disconfirmation of expectations in which pre- and post-trial information regarding the same dimensions must be integrated.

FIGURE 3

CHANGES IN BELIEF STRENGTH FOR BELIEF-EXPECTANCIES REGARDING RELATIVE EXPENSIVENESS



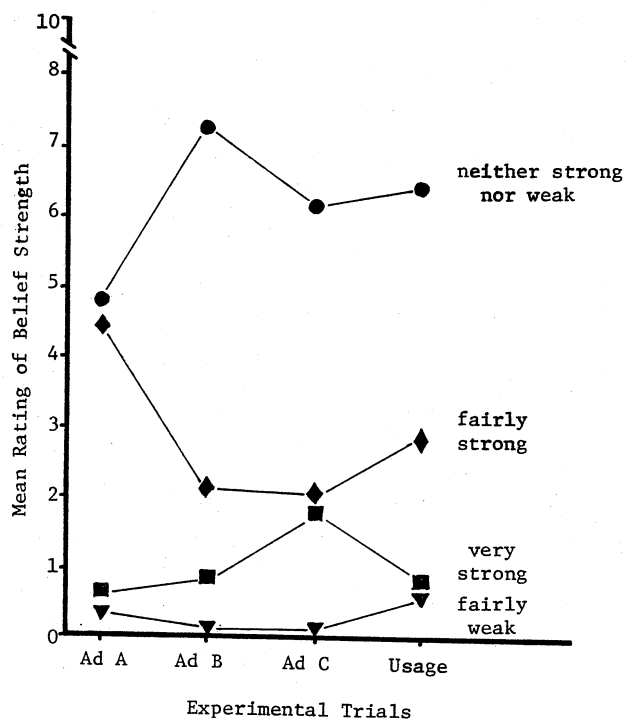
Briefly, two other theories may be useful for developing further research regarding the disconfirmation of consumer expectations. Helson's (1964) adaptation level theory has been applied in a price perception context (Nwokoye, 1975) and may be relevant to understanding belief-expectation development. Comparison level theory (cf. Thibaut and Kelley, 1959) has been used by Swan (1972) in a consumer expectation study focusing on information search processes.

Finally, a cognitive phenomena involving several interesting theoretical belief-related issues warrants brief discussion. If it is assumed that belief dimensions are somehow interrelated in a cognitive structure, the creation of a belief-expectancy regarding one product attribute may be considered to have "spill-over" effects on belief-expectancies about other attribute dimensions (called inferential belief formation by Fishbein and Ajzen, 1975). Unfortunately, few consumer behaviorists have considered cognitive spill-over effects in their research (for an exception, see Lutz, 1975).

The present study allowed for the determination of cognitive spill-over effects from the bitterness belief-expectancy manipulations including measures of four other belief dimensions about ground coffee (see Table 2). Due to space constraints, only two of these belief vectors are examined in this paper. Figures 3 and 4 present the changes in belief-expectancies for the "level of expense" and "strength of flavor" attributes over the four experimental trials. Both figures indicate that the three expectancy-creating messages regarding "Coffee's" bitterness had some spill-over effects in terms of creating other belief expectancy elements in respondents' cognitive structure. Note that the belief strength ratings for "level of expense" and "strength of flavor", even though not directly manipu-

FIGURE 4

CHANGES IN BELIEF STRENGTH FOR BELIEF-EXPECTANCIES REGARDING STRENGTH OF FLAVOR



lated, generally seem to approach a relatively stable or asymptotic pretrial level after the second message exposure. This pattern is similar to that found for the manipulated bitterness belief vector.

The cognitive process by which related belief-expectancies are acquired, even though information specific to those dimensions has not been processed, is one of great interest. Cognitive spill-over effects have obvious methodological and theoretical implications for those interested in attitude change, advertising effectiveness, or tests of various models of cognitive structure such as expectancy-value models. Although we will not speculate regarding the form of the internal information integration (or generation?) processes underlying these observed spill-over phenomena, several variables seem likely to be positively related to the amount of spill-over obtained. Among these candidates for incorporation into future research are: the amount of experience with a product, the complexity of the cognitive structure (e.g., number of belief dimensions) for a product, the level of confidence in one's belief-expectancy ratings (perhaps indicating the ambiguity or instability of certain belief dimensions), and the centrality of the manipulated belief in one's cognitive structure.

#### Conclusions

The conceptualization of consumer expectancies as specific belief elements within a cognitive structure appears to have been successful. Moreover, these belief-expectancies seem measurable with sufficient precision using the "vector model" methodology that relatively small and subtle belief changes can be identified.

Considering consumer product expectancies as a type of belief has obvious advantages, foremost among which is the ability to rely on established and proposed theoretical linkages between beliefs and other cognitive constructs. This facilitates theorizing about expectancy creation and disconfirmation, as well as developing of testable models of expectancy-related phenomena. Although several potentially useful theoretical perspectives have been discussed, most with a cognitive structure orientation, the task of developing clear-cut tests of competing theories will clearly be difficult. Presently, the major conceptual perspectives seem to be dissonance and assimilation-contrast theories. It may prove useful to base future investigations upon notions derived from information processing research.

#### References

- Olli T. Ahtola, "The Vector Model of Preferences: An Alternative to the Fishbein Model," *Journal of Marketing Research*, 12 (February 1975), 52-9.
- Ralph E. Anderson, "Consumer Dissatisfaction: The Effect of Disconfirmed Expectancy on Perceived Product Performance," *Journal of Marketing Research*, 10 (February 1973), 38-44.
- Ralph E. Anderson and Joseph F. Hair, "Consumerism, Consumer Expectations, and Perceived Product Performance," *Proceedings, Third Annual Conference of the Association for Consumer Research* (1972), 812-815.
- James R. Bettman, Noel Capon, and Richard J. Lutz, "Cognitive Algebra in Multi-Attribute Models," *Journal of Marketing Research*, 12 (May 1975), 151-164.
- Stewart W. Bither, and Stephen J. Miller, "A Cognitive Theory View of Brand Preference," *Proceedings, Fall Conference of the American Marketing Association* (1969), 280-286.

- Donald T. Campbell, and Julian C. Stanley, Experimental and Quasi-Experimental Designs for Research (Chicago: Rand McNally, 1963).
- Richard N. Cardozo, "An Experimental Study of Customer Effort, Expectation and Satisfaction," Journal of Marketing Research, 2 (August 1965), 244-249.
- Joel B. Cohen, and Marvin E. Goldberg, "The Dissonance Model in Post-Decision Product Evaluation," Journal of Marketing Research, 7 (August 1970), 315-21.
- Joel B. Cohen, Martin Fishbein, and Olli T. Ahtola, "The Nature of Uses of Expectancy-Value Models in Consumer Attitude Research," Journal of Marketing Research, 9 (November 1972), 456-60.
- Martin Fishbein, "An Investigation of the Relationships Between Beliefs About an Object and the Attitude Toward that Object," Human Relations, 16 (August 1963), 233-240.
- Martin Fishbein, "A Consideration of Beliefs and Their Role in Attitude Measurement," in Martin Fishbein (ed.), Readings in Attitude Theory and Measurement. (New York: Wiley, 1967), 257-66.
- Martin Fishbein and Bertram H. Raven, "The AB Scales: An Operational Definition of Beliefs and Attitudes," Human Relations, 15 (1962), 35-44.
- Martin Fishbein and Icek Ajzen, Belief, Attitude, Intention and Behavior (Reading, Massachusetts: Addison-Wesley Publishing Company, 1975).
- Fritz Heider, "Attitudes and Cognitive Organization," Journal of Psychology, 21 (January 1946), 107-12.
- Harry Helson, "Current Trends and Issues in Adaption-Level Theory," American Psychologist, 19 (1964), 26-38.
- George Katona, "Psychology and Consumer Economics," Journal of Consumer Research, 1 (June 1974), 1-8.
- Eugene J. Kelley, and Rusty Scheewe, "Buyer Behavior in a Stagflation/Shortages Economy," Journal of Marketing, 39 (April 1975), 44-50.
- Richard J. Lutz, "Changing Brand Attitudes Through Modification of Cognitive Structure," Journal of Consumer Research, 1 (March 1975), 49-59.
- Michael B. Mazis, Olli T. Ahtola and R. Eugene Klippel, "A Comparison of Four Multi-Attribute Models in the Prediction of Consumer Attitudes," Journal of Consumer Research, 2 (June 1975), 38-52.
- Theodore M Newcomb, "An Approach to the Study of Communication Acts," Psychological Review, 60 (1953), 393-404.
- Nonyelu Nwokoye, "An Experimental Inquiry Into the Effects of Parameters of Price Structure on Buyers' Price Judgments: A Condensation of a Doctoral Dissertation," School of Business Administration, University of Massachusetts, Amherst, Working Paper #75-12 (May 1975).
- Richard N. Olshavsky, and John A. Miller, "Consumer Expectations, Product Performance, and Perceived Product Quality," Journal of Marketing Research, 9 (February 1972), 19-21.
- Jerry C. Olson, "Cue Properties of Price: Literature Review and Theoretical Consideration," Working Series in Marketing Research, College of Business Administration, Pennsylvania State University (1974), Paper #20.
- Jerry C. Olson, and Andrew A. Mitchell, "The Process of Attitude Acquisition: The Value of Developmental Approach to Consumer Attitude Research," in M. J. Schlinger (ed.), Advances in Consumer Research: Volume II. (Chicago: Association for Consumer Research, 1974), 249-264.
- Jerry C. Olson, Daniel R. Toy, and Philip Dover, "Consumer Expectations: Literature Review and Theoretical Considerations," Working Series in Marketing Research, College of Business Administration, Pennsylvania State University (1975), Paper #36.
- Charles E. Osgood, and Percy H. Tannenbaum, "The Principle of Congruity in the Prediction of Attitude Change," Psychological Review, 62 (1955), 42-55.
- Milton J. Rosenberg, "Cognitive Structure and Attitude Affect," Journal of Abnormal and Social Psychology, 53 (November 1956), 367-72.
- Alan G. Sawyer, "Demand Artifacts in Laboratory Research on Consumer Behavior," Journal of Consumer Research, 1 (March 1975), 20-29.
- Muzafer Sherif and Carl I. Hovland, Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change (New Haven, Connecticut: Yale University Press, 1961).
- Jagdish N. Sheth, "A Field Study of Attitude Structure and Attitude Behavior Relationship," in J. N. Sheth (ed.), Models of Buyer Behavior (New York: Harper & Row, 1974).
- John E. Swan, "Search Behavior Related to Expectations Concerning Brand Performance," Journal of Applied Psychology, 56 (August 1972), 332-335.
- J. W. Thibaut, and H. H. Kelley, The Social Psychology of Groups (New York: Wiley, 1959).
- Edward Chace Tolman, Purposive Behavior in Animals and Men (New York: Appleton-Century-Crofts, 1932).
- William L. Wilkie, and Edgar A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10 (November 1973), 428-41.
- David T. Wilson, and Philip Dover, "A Test of the Repertory Grid Technique as a Means of Developing Attributes for Use in Choice Models," Working Series in Marketing Research, College of Business Administration, Pennsylvania State University (1975), Paper #33.
- Arch G. Woodside, "Positive Disconfirmation of Expectation and the Effect of Effort on Evaluation," Proceedings, 80th Annual Convention, APA (1972), 743-744.

AN INVESTIGATION INTO THE DIFFERENTIAL EFFECTS  
OF CAUSALLY SIMPLE AND COMPLEX ATTRIBUTIONS

Richard Mizerski, Arizona State University

Abstract

This study examines a proposed measure of consumer causal attribution that is based upon the stochastic assumptions of Information Theory. Given the same product information, significant differences were found between causally simple and complex subjects in their attributions, beliefs, and affect toward the stimulus products.

Attribution in Information Processing

The concept of causal attribution, and the study of its impact on the interpretation of information, is a relatively new and expanding area of study by social psychologists. Since information-processing is a critical area in many phases of marketing, it was only natural that researchers would investigate the feasibility of applying these same concepts to help explain consumer processing of product or service information in the marketplace. The few marketing studies that have been conducted (e.g., Mizerski, 1974a; Robertson and Rossiter, 1974; Settle, 1973; Settle and Golden, 1974) show that the concepts can be generalized to marketing, and may provide new perspectives for understanding the consumer. Nonetheless, research in both marketing and social psychology has been severely limited in examining the individual's total attribution process. The first problem is that the causal attributions in these studies are seldom measured directly. Instead, they are inferred using such measures as confidence and credibility. This procedure follows from Jones and Davis (1965), who suggested an isomorphic relationship between confidence and extremity of the attribution. However, research by Newton (1973) and Mizerski (1974b) seriously questions the validity of this approach. Second, present attribution measures do not allow for an examination of the number of perceived causes, nor the allocation of attribution among each cause. Since consumers generally feel that information from or about the marketplace is the result of a mix of internal and various external causes (Mizerski, 1974a; 1974b), improvements in the elicitation and measurement of this mix should provide for more precise tests and applications of attribution principles. Toward that end, this paper presents a new measure of causal attribution, causal complexity, that is based upon the stochastic assumptions of Information Theory.

Attribution and the Concept of Causal Complexity

Attribution Theory studies the processes whereby people make causal explanations about the information they receive. Then, based upon the causal attribution chosen, it attempts to predict how those individuals will make inferences about their environment. In terms of processing information in the marketplace, consumers can contribute product information to either an internal, external, or a mix of both types of causes. For example, suppose that a consumer received information that a particular brand or model of automobile delivered good gas mileage. If the receiver believed that the information was the result ("is caused by") the information source actually getting good gas economy and then telling the receiver about it, he would attribute the information to an internal cause--a cause internal to the stimulus object. In many cases, however, the consumer may have some doubt about whether the informa-

tion was "caused" by actual product performance. The receiver may suspect that other causes such as the source's conflict of interest (e.g., the source is selling his own automobile), or some other individual bias (e.g., the source always prefers to speak favorably about the brand) prompted the information. In these two latter situations, the receiver attributes the information to "external causes" and generally finds the information much less suitable for making inferences about the product.

However, the degree to which a consumer makes an internal attribution is not necessarily the only factor of interest. Consumer perception of what is suitable/useful information is largely determined "... by the configuration of factors that are plausible causes for that information" (Kelley, 1973, p. 108). In other words, the type and mix of external causes the consumer believes exists can also tell us a great deal about how he or she will interpret product information.

Unfortunately, present scaling techniques can make only very simple distinctions between causal attributions. As an example of this limitation, suppose there are four plausible reasons (causes) for a source to transmit product information (Figure 1). Cause #1 is the internal

Figure 1

Hypothetical Causal Allocation of Product Information

Cause	Receiver		
	A	B	C
Internal 1	20%	20%	20%
2	80%		20%
External 3		40%	20%
4			20%
5		40%	20%

attribution (the product itself "caused" the information transmission), while causes 2 through 5 are alternate, non-product causal attributions (e.g., source bias, peer pressure to say something favorable or unfavorable about the brand). Note that all three hypothetical receivers allocate 20 percent of the cause to the product (internal cause), with the remainder of the allocation going to various external causes. Present techniques will examine only part of each receivers' attribution domain by measuring the degree of internal attribution (A=20 percent, B=20 percent, and C=20 percent) or evaluate attributions made between an internal and a surrogate for all external causes (A=20 percent internal, 80 percent external; B=20 percent internal, 80 percent external; and C=20 percent internal, 80 percent external). While this approach may be effective for some types of research, it tends to view the causal attribution of receivers A, B and C as identical. It would seem, however, that there are very different attributional processes involved, and that the "complexity" of each subject's causal allocation (i.e., in terms of number dimensions or causes the attributors employ) may be an important factor that affects how he or she will react to product information (Kelley, 1972).

Figure 2 provides another example of the attribution of information by a number of receivers. Receiver A attributes the information to only one cause while receivers

Figure 2

Hypothetical Causal Allocation of Product Information

Cause	Receiver				
	A	B	C	D	E
1	100%	20%	96%	50%	50%
2		20%	1%	50%	25%
3		20%	1%		15%
4		20%	1%		10%
5		20%	1%		

B and C both perceive four plausible causes for the same effect (the product information). Given information about (1) the number of perceived causes; (2) the allocation of causality; and (3) how a specific cause affects the interpretation of an effect, a marketer could predict receiver A's interpretation and use of the data with more certainty than for either receiver B or C. While receiver C sees more plausible causes than A, he attributes the great bulk of causation to the same factor (#1), and should form very similar product perceptions with the information. Therefore, the marketer would be only slightly less certain about receiver C's information processing. Receiver B, however, feels that each of the causes are equally responsible for the information. This causal allocation provides, in effect, equally weighted explanations about how the consumer will react, and makes the marketer least certain about predicting B's resultant interpretation and use of the product information. It should be noted that this uncertainty is also operating within the attributor himself. The receiver finds situations where many causes are equally plausible more difficult to interpret than when an effect can be attributed to one or a few causes (Jones and Davis, 1965). Therefore, the degree of uncertainty for both the attributor and the marketing observer is a direct function of the dispersion of attribution among the set of plausible causes.

Measuring Causal Complexity

How may one get an objective overall measure of an individual's causal complexity? From the example in Figure 2, it should be apparent that both the number of perceived causes as well as the allocation of cause must be included, since neither completely describes the uncertainty in a receiver's causal array. Fortunately, the question of how to gauge the complexity of an array of elements has already presented itself in such diverse contexts as psychology, statistics, and communications, and has spurred development of an index of uncertainty called "H" or entropy. This technique was initially proposed by Shannon and Weaver (1949) to judge the amount of uncertainty<sup>1</sup> in an array of symbols by measuring the rectilinearity of their distribution. If one considers a subject's decisions on the number and allocations of cause to be a causal array the applica-

<sup>1</sup>To the information theorists ". . . uncertainty and information are interchangeable commodities . . . (with) the amount of information in an event exactly equal to the amount of uncertainty residing in that event before its occurrence, and this uncertainty, in turn, is a direct function of the number of possible events that could have occurred" (Bieri, et. al., 1966, p. 51).

tion of H to measure causal complexity is rather straight-forward. Array complexity is calculated by accounting for the number of distinctions made about a domain in terms of the grouping system used by the subject. The following is the formula for H:

$$H = - \sum_{i=1}^K P_i \log_2 P_i$$

where: H is the individual subject's causal complexity score;

$P_i$  is the allocation to cause i; and

K is the number of perceived causes.

Applying this formula to the attributions in Figure 2 would designate receiver A as the most causally simple (least uncertainty) and receiver B as the most complex (most uncertainty). The complexity continuum would read, from causally simple to complex; A, C, D, E, and B. Complexity, and thus the highest H score, is greatest with equal allocation of causal attribution among the largest number of perceived causes; while the least complexity would be reflected in the tendency to attribute to one cause.

Suggested Differences Due to Causal Complexity

Given a measure of causal complexity, previous research suggests that a number of differences should exist between causally simple and complex individuals.<sup>2</sup> First, the more complex the causal domain, the less confident the individual should be in his or her attribution about the information. This concept forms the basis for the proposed measure, that individuals are less certain/confident in a "broad distribution" (causal allocation) than they are in a "sharply peaked one" (Jaynes, 1957; Jones and David, 1965). Second, given credible product information, causally simple subjects should tend to be more extreme in their causal attributions.<sup>3</sup> This follows from the "discounting principle" that suggests "the role of a given cause in producing a given effect is discounted if other plausible causes are also present" (Kelley, 1973, p. 113). Without information to the contrary, however, the internal attribution tends to be expected (that the information about the product was caused by the product). Therefore, a large number of external attributions should discount (in the consumer's mind) the possibility of an internal cause.

Third, given the same product information, causally simple individuals should form more extreme beliefs about the product. In an approach consistent with other researchers, Fishbein (1965, p. 107) has stated that ". . . a belief about an object may be defined as the probability or improbability that a particular relationship exists between the object of belief and some other object, concept, value, or goal." This definition of a belief appears to be one result of an attribution (Ajzen and Fishbein, 1975; Mizerski, 1974a). Kelley (1973,

<sup>2</sup>This categorization should not be interpreted as a personality variable or anything beyond a situation specific attribution decision at this point in the constructs development. Future research must establish the extent of the generalizability across product and information processing situations.

<sup>3</sup>The type of attribution (internal versus external) depends on a number of consumer and situational factors (e.g., the source of the information).



p. 107) notes that attribution is a part of the process ". . . by which man 'knows' his world, (and) has a sense that his beliefs and judgments are veridical." It would seem to follow that if an individual made a strong attribution about information concerning product characteristics to an internal cause, this would manifest itself in a belief that a relationship existed between the product and those characteristics. In Fishbein terminology, the internal attribution attaches a higher probability that the information (e.g., product attribute ratings) was related to the product. The stronger the internal attribution, the more extreme the belief. If causal complexity follows the discounting principle, causally simple individuals would make stronger internal attributions that should result in more extreme beliefs. If the simple subjects have a greater degree of confidence in their attributions (the first proposed difference), that should further augment a differential in belief strength.

Finally, causally simple and complex subjects should differ in terms of the extremity of their attitude toward a product, given the same stimulus information. A substantial amount of research on attitude models (see Bass and Wilkie, 1973, for a review) suggests that beliefs, defined in the Fishbein sense, are one of the major components in forming an individual's attitude toward a product. One suggested algebraic relationship between the two (Fishbein, 1965, p. 117) is the following:

$$A_o = \sum_{i=1}^N B_i a_i$$

where:  $A_o$  = measure of affect;

$B_i$  = the strength of belief  $i$  about the attitude of object  $o$ , that is the probability or improbability that  $o$  is related to some other object  $x_i$ ;

$a_i$  = the evaluative aspect of  $B_i$ , that is, the evaluation of  $x_i$ --its goodness or badness;

$N$  = the number of beliefs.

While other expectancy-value models differ somewhat from Fishbein's, each uses some measure of instrumentality or belief strength as a basis for predicting an attitude. If causally simple subjects form more extreme beliefs, the differential in extremity should manifest itself in a more extreme affect toward the product.

#### Hypotheses

In order to test for the suggested difference between causally simple and complex individuals, the following hypotheses were developed:

Given the same stimulus product information;

1. Causally simple subjects will tend to be more confident in their causal allocations.
2. Causally simple subjects will tend to be more internal in their causal attributions.
3. Causally simple subjects will tend to form more extreme beliefs about a product.
4. Causally simple subjects will tend to produce more extreme affect toward the product.

#### Data Collection

An experiment was performed on 300 upper division undergraduate and graduate students at the University of Florida College of Business Administration.<sup>4</sup> The subjects were told that the experiment involved testing how people evaluate, and are affected by, the opinions of others. They then received evaluative information on three salient attributes (chosen by the Fishbein and Raven, 1962, procedure) of either a fictitious automobile or motion picture. Attributes used were maintenance costs, comfort, and gas mileage for the automobile; and acting, plot, and photography for the movie. The information was presented in the form of personal ratings that were supposedly made by another student, who was randomly chosen to test and evaluate the product. Two of the three attributes were given a neutral rating, with only one attribute rated either favorably or unfavorably.<sup>5</sup> The modifiers used for rating were scaled for equal polarity and opposite affective meaning by the methods used in Myers and Warner (1968). Subjects were randomly assigned to the attribute (3 levels) x information (2 levels) x product (2 levels) treatments.

#### The Measure of Causal Complexity

Following the information treatment, the subjects were asked to allocate the cause for the bogus "rater's" opinion (the information treatment) with the following question concerning attribution complexity:

How much do you feel that each of the following contributed to the student's opinion about gas mileage? (note: any of the following could account for 0% to 100% of the opinion)

- |  |         |
|--|---------|
| 1. The automobile itself. . . . .  | _____ % |
| 2. The influence of other peoples' opinions . . . . .  | _____ % |
| 3. An effort to please or antagonize the viewer . . . . .                                      | _____ % |
| 4. The personality of the tester (natural tendencies to be critical or complimentary). . . . . | _____ % |
| 5. A general bias for or against automobiles or the brand . . . . .                            | _____ % |
| 6. Other reasons (if any)? _____   | _____ % |
| TOTAL  | 100 %   |

The external causes (numbers 2 through 6) were determined in two pretest group sessions with subjects similar to those in the main experiment. Each subject's causal complexity score was produced by applying the H transformation to their responses about causal allocation. The respondents were then ordered by complexity, with analyses performed on subjects in the upper and lower quartiles (most and least causally complex). A seven point scale ranging from "no confidence" (#1) to "complete confidence" (#7), recorded the subjects' confidence in the causal allocation for each treated attribute.

<sup>4</sup>For the analyses, 30 subjects were deleted for missing data and for improper use of a scale.

<sup>5</sup>Two information treatments were used in order to conduct another experiment that evaluated the disproportionate influence of unfavorable information (Mizerski, 1974a; 1974b).

The Measure of Internal Attribution

A second question was developed to establish an independent measure of the subject's degree of attribution to an internal cause.<sup>6</sup>

To what extent do you feel that other reasons--reasons having nothing to do with the automobile tested--influenced the student's opinion about. . .

	other reasons had no effect on the opinion				other reasons were the only cause for the opinion			
gas mileage	1	2	3	4	5	6	7	

The smaller the scale value chosen, the more internal ("caused" by the product only) the attribution.

The Measure of Belief

A third question elicited the subjects' belief about the relationship between the treated attribute and the product, and randomly appeared either before or after the two attribution questions (the order of the attribution questions were also randomized). The format for the belief scale follows that developed by Fishbein and Raven (1962), and has been used in a number of marketing studies (e.g., Mazis and Klippel, 1973; and Mizerski, 1974a).

How likely is it that this automobile . . .

unlikely									likely
has low gas consumption	-4	-3	-2	-1	0	+1	+2	+3	+4

The Measure of Attitude

Finally, subjects were asked a question about their overall affect toward the stimulus product.

How much would the automobile tested appeal to you?

Extremely Low Appeal	Mildly Low Appeal	Mildly Low Appeal	Neutral	Mildly High Appeal	Extremely High Appeal	Extremely High Appeal
1	2	3	4	5	6	7

Results

Mean scores for the measure of the subjects' confidence in their attribution allocation are presented in Table 1. As hypothesized, causally simple subjects, those who perceived relatively few causes for the stimulus information, were more confident ( $\bar{X}_S = 5.16$ ) than complex subjects ( $\bar{X}_C = 4.54$ ) in their causal allocation.

<sup>6</sup>See Mizerski (1974b) for a validation and other uses for this scale.

Table 1

	Mean Confidence Scores					
	Simple (5.16)			Complex (4.54)		
Information	Auto	Movie	Combined	Auto	Movie	Combined
Favorable	5.40	5.15	5.28	4.25	4.37	4.31
Unfavorable	5.31	4.79	5.05	4.42	5.12	4.77
Combined	5.36	5.06		4.34	4.75	

In order to detect if these differences were significant, an unweighted means analysis of variance (ANOVA) was performed on the data (Table 2). The ANOVA revealed a main effect of complexity ( $F = 6.73, p < .01$ ) which supports the first hypothesis.

Table 2

Analysis of Variance:  
Confidence in the Causal Allocation

Source	d.f.	MS	F
Causal Complexity (C)	1	12.47	6.73 <sup>a</sup>
Product (P)	1	0.00	0.00
Information treatment (I)	1	0.43	0.23
CXP	1	5.08	2.74
CXI	1	3.82	2.06
PXI	1	0.19	0.10
CXPXI	1	1.46	0.78
Error	126	1.85	

ap  $\leq$  .01

The second hypothesis, that causally simple subjects would be more internal in their attribution, was also supported. Mean scores for the independent measure of internal attribution are presented in Table 3.

Table 3

	Mean Internal Attribution Scores					
	Simple (2.92)			Complex (3.76)		
Information	Auto	Movie	Combined	Auto	Movie	Combined
Favorable	2.70	3.35	3.03	3.83	3.95	3.67
Unfavorable	2.92	2.71	2.82	4.25	3.00	3.63
Combined	2.81	3.03		4.05	3.48	

Since a low mean score reflects an internal attribution, the difference between causally simple and complex subjects are in the predicted direction. An ANOVA on the mean scores (Table 4) reveals a main effect of causal

complexity ( $F = 8.55, p < .01$ ), and an interaction of product and information ( $F = 3.79, p < .05$ ).

Table 4  
Analysis of Variance:  
Internal Cause Attribution

Source	d.f.	MS	F
Causal Complexity (C)	1	22.72	8.55 <sup>a</sup>
Product (P)	1	1.01	0.38
Information treatment (I)	1	1.75	0.66
CXP	1	5.10	1.92
CXI	1	0.02	0.01
PXI	1	10.09	3.79 <sup>b</sup>
CXPXI	1	0.54	0.20
Error	126	2.66	

<sup>a</sup> $p < .01$

<sup>b</sup> $p < .05$

The main effect of causal complexity shows that the internal attribution of the causally simple subjects were significantly larger than those of the complex group, and thus supports the second hypothesis. The interaction of product and information was independent of the subject's complexity and does not affect the hypothesis (see Mizerski, 1974b, for a discussion of this interaction).

Analyses of hypothesis three, that causally simple and complex subjects would differ in terms of their extremity of belief given the same stimulus information, required some initial data transformation. Since differences in absolute belief strength were of interest, the sign of the beliefs formed in the unfavorable information treatment (usually 0 to -4 on the -4 to +4 scale) were reflected.<sup>7</sup> The mean absolute belief scores for all subjects are presented in Table 5. As predicted,

Table 5

Mean Belief Scores

	Simple (2.81)			Complex (2.34)		
	Auto	Movie	Combined	Auto	Movie	Combined
Information						
Favorable	3.30	2.50	2.90	1.92	2.68	2.30
Unfavorable	3.08	2.36	2.72	2.58	2.18	2.38
Combined	3.19	2.43		2.25	2.43	

<sup>7</sup>The term "reflected" refers to the procedure of changing the sign of the value (e.g., -4 becomes +4). The method used in this study was to multiply the belief score of subjects in the unfavorable information treatment by a -1.

the overall mean belief score for the simple subjects ( $\bar{X}_S = 2.81$ ) was more extreme than the causally complex group's ( $\bar{X}_C = 2.34$ ) beliefs. The high and low complexity groups also differed in terms of the extremity of their beliefs formed about the treated automobile attribute, with simple subjects forming much stronger beliefs ( $\bar{X}_S = 3.19$ ) than the complex individuals ( $\bar{X}_C = 2.25$ ). Applying an ANOVA to these scores (Table 6 reveals both a significant main effect of causal complexity ( $F = 4.63, p < .05$ ) and an interaction of the subject's complexity and the product. In order to detect the source of the complexity x product interactions, a Newman-Keuls Multiple range test was performed (Table 7). The results substantiate that there were significant differences between the simple and complex subjects' beliefs about the treated automobile attribute. This supports the hypothesis that causally simple subjects will form more extreme beliefs, and further notes the significant effect that the nature of the product can have on the attribution/belief formation process.

TABLE 6

Analysis of Variance:  
Belief Scores

Source	d.f.	MS	F
Causal Complexity (C)	1	7.11	4.63 <sup>a</sup>
Product (P)	1	2.69	1.75
Information treatment	1	0.09	0.06
CXP	1	7.17	4.67 <sup>a</sup>
CXI	1	0.55	0.36
PXI	1	2.40	1.56
CXPXI	1	3.15	2.05
Error	126	1.54	

<sup>a</sup> $p < .05$

TABLE 7

Results of a Newman-Keuls Multiple Range Test

Causal Complexity	Comparison Groups			
	Complex	Simple	Complex	Simple
Product	Auto	Movie	Movie	Auto
Means	2.25	2.43	2.43	3.19

-----Not significantly different at  $p < .05$

The final hypothesis predicted that, given the same information, causally simple subjects would form more extreme attitudes or affect toward the product. This hypothesis was based largely upon the suggested existence of differences between the two causal styles in attribution confidence, perceived internal causation, and extremity of beliefs. Since these suggested differences did occur, support for the last hypothesis was anticipated.

Since approximately half of the subjects received unfavorable information about one or the other of the products, the original affect scale (#1: extremely low appeal to #7: extremely high appeal) could not be used to test for differences in extremity. Therefore, each score was transformed to a -3/+3 scale value, and affect ratings of those subjects who received unfavorable information were reflected. Mean affect scores are reported in Table 8.

TABLE 8

Mean Affect Scores

	Simple (0.91)			Complex (0.44)		
	Auto	Movie	Combined	Auto	Movie	Combined
Favorable	0.70	0.50	0.60	-0.17	0.36	0.58
Unfavorable	0.92	1.50	1.21	0.74	0.82	0.78
Combined	0.81	1.00		0.29	0.62	

As hypothesized, causally simple subjects tended to form a more extreme attitude toward both products ( $\bar{X}_S = 0.91$ ,  $\bar{X}_C = 0.44$ ). Analysis of variance on the affect scores (Table 9) show main effects for both causal complexity ( $F = 5.24$ ,  $p \leq .025$ ) and the information treatment ( $F = 10.09$ ,  $p \leq .01$ ). While the latter results do not constitute a test of the hypothesis, they provide additional evidence that unfavorable information is more influential, and that this disproportionate influence seems to be linked to a difference in the attribution process (for an extensive discussion of this phenomenon, see Mizerski, 1974a and 1974b). The significant differences in affect extremity, with causally simple subjects forming more extreme attitudes, supports the last hypothesis.

TABLE 9

Analysis of Variance:  
Extremity of Attitude/Affect

Source	d.f.	MS	F
Causal Complexity (C)	1	6.99	5.24 <sup>b</sup>
Product (P)	1	2.01	1.51
Information treatment (I)	1	13.44	10.09 <sup>a</sup>
CXP	1	.12	.09
CXI	1	.04	.03
PXI	1	.22	.16
CXPXI	1	3.03	2.27
Error	126	1.33	

<sup>a</sup><sub>p</sub> ≤ .01

<sup>b</sup><sub>p</sub> ≤ .025

## DISCUSSION

Results of the experiment support the validity of causal complexity, as well as suggest a number of potentially important cognitive differences between the two causal types. Causally simple subjects tended to be more confident in their attributions, and more apt to attribute the information to an internal cause (the product itself). They also formed more extreme beliefs and attitudes about the stimulus product; the first time attribution has been empirically linked to these latter two measures. It should also be noted that none of the results, except for the product x information interaction for beliefs, were significantly affected by the product on which information was presented. While future research should gauge the extent to which these results can be generalized, the two products used represented very different degrees of social visibility and desirability (in many ways representing a good versus a service), and suggests that the subjects' causal complexity may be rather insensitive to the specific product being evaluated.

A particularly difficult question is establishing where causal complexity fits into the sequence of consumer information-processing, since a number of relationships are plausible. Causal complexity may operate as a personality trait, remaining consistent across many product classes. This relationship required that the subjects complexity scores would have to be largely independent of the product information. That appears doubtful in this experiment since the subjects' complexity scores were significantly affected by one variable in the experimental treatment. An ANOVA on these scores (Table 10) shows a main effect of product attribute ( $F = 8.87$ ,  $p \leq .001$ ); yet it is interesting to

TABLE 10

Analysis of Variance: Causal Complexity

Source	d.f.	MS	F
Product (P)	1	.075	.874
Attributes (A)	4	.764	8.87 <sup>a</sup>
Information treatment (I)	1	.028	.33
PXI	1	.016	.184
AXI	4	.047	.55
Error	258	.086	

<sup>a</sup><sub>p</sub> ≤ .001

note that neither the product nor the information treatment (favorable or unfavorable product information) prompted significant differences in subjects' complexity.

Mizerski (1974b) found similar main effects of product attribute related to differences in subjects' attribution of internal causation about product information. It was suggested that the different degree of objectivity of judgment between the attributes caused this effect. In other words, the more an attribute lends itself to a subjective evaluation (e.g., styling vs. gas mileage), the greater the possibility that the subjects will perceive external causes for that information. In terms of this study, the subjectivity of an attribute's evaluation may trigger the subjects'

causal complexity, seemingly a very situation specific reaction.

If causal complexity is situational, it could act as a mediating or contributing factor in the determination of the consumer's beliefs and attitude toward a product. Complexity's differential response (compared to the belief and affect scores) to the experimental treatments, however, tends to cast doubt that it operates as the sole criterion. Rather, causal complexity may be formed by the information, then the resultant complexity, situational variables, and the information itself, all prompt relevant beliefs and attitudes.

Of course, another alternative explanation is that the product information prompts the subjects to make causal attribution, which in turn establishes their causal complexity. While the attribution literature suggests that causal perception leads to attribution, this explanation cannot be ruled out at this time (if ever). The determination of causal complexity's place in consumer information-processing will require a measure of complexity independent of the treatments.

Future research could also examine the generalizability of this study's results by introducing such factors as the subjects' interest and experience with the product and the impact of various sources of information (e.g., consumer groups, trade associations, government agencies) on the complexity of the attribution. It should also be noted that the size of the subjects' causal domain was in some ways fixed in this study. It would be interesting to see if the use of a relative rather than an absolute measure of complexity would provide a better measure, outweighing the operational problems involved.<sup>8</sup>

The data also suggest plausible ties with another area of behavior, cognitive complexity. While the two constructs theoretically represent different psychological processes, they show similar results in a number of areas (see Streufert, 1972, for a review). Both causally and cognitively simple individuals tend to perceive more internal attributions in certain situations (Harvey and Ware, 1967), as well as form more extreme affect toward a stimulus object or person (Campbell, 1960). Other aspects, such as a differential in the type and extent of information search the two cognitive styles exhibit, may also be applicable to causal complexity. The cognitive complexity literature is rich in empirical research and could prove helpful in further explaining and applying the concept of causal complexity.

<sup>8</sup>The H formula can be made a relative measure of causal complexity (R) by the following modifications.

$$R = \frac{H}{H_{\max}}$$

Where  $H_{\max}$  is a calculation for each subject when subjects perceive different numbers of plausible causal categories. R is calculated by dividing the subject's H by the maximum H possible with that number of categories.

By using a relative, rather than an absolute measure of causal complexity, the experimenter could elicit the subject's causal domain with a free-response questioning technique. This would overcome the problem of possible confining some subjects as to the size of their relevant causal choices. The trade-off is that difficult, if not impossible, subjective evaluations must be made determining if each subject's causal categories are independent.

The study's results suggest several possible marketing applications. The measure may enable better prediction of consumer response to product information. Kelley (1972, 1973) has developed a framework called a causal schema which is a conceptual representation of how two or more causal factors (e.g., internal vs. external) interact in relation to a particular type of effect (e.g., transmitted product information). He suggests that individuals learn through experience to link effects with possible causes, thus building a ". . . repertoire of abstract ideas (schemata) about the interaction of causal factors" (Kelley, 1972, p. 152). These schemata provide the individual with an economical and fast attribution framework for the fitting of partial information so that reasonably good causal inferences can be drawn. If the marketer knows that a specific schema represents a particular causal problem, he can then make a reasonably good prediction as to how the consumer will interpret the information provided. Although Kelley has developed rather elaborate and complex schemata that integrate many types of causal attributions, the necessary behavioral information required to test and implement these constructs has not been available. Thus, present applications have been limited to representing rather simple attribution decisions (main effects, not causal interactions) which severely limits predictability in more complex market situations. Kelley notes that an important unanswered question concerning the ". . . complexity, in terms of the dimensions or causes . . . attributors can and do employ in their schemata." Without this knowledge we are ". . . uncertain about the kind of causal schema being used" (Kelley, 1972, p. 156). Causal complexity is a more inclusive measure of the individual's causal domain, and may provide the necessary information about the consumer's complexity, and ultimately the information-processing strategy he or she is using.

#### References

- Icek Ajzen, and Martin Fishbein, "A Bayesian Analysis of Attribution Processes," Psychological Bulletin, 2 (1975), 261-277.
- Frank M. Bass and William L. Wilkie, "A Comparative Analysis of Attitudinal Predictions of Brand Preference," Journal of Marketing Research, 10 (1973), 262-269.
- S. Bieri, A. L. Atkins, S. Briar, R. L. Leaman, H. Miller and T. Tripodi, Clinical and Social Judgment (New York: Wiley, 1966).
- V. N. Campbell, Assumed Similarity, Perceived Sociometric Balance and Social Influence, unpublished doctoral dissertation, University of Colorado, 1960.
- Martin Fishbein, "A Consideration of Beliefs, Attitudes, and Their Relationship," in Ivan D. Steiner and Martin Fishbein, eds., Current Studies in Social Psychology, (New York: Holt, Rinehart, and Winston, Inc., 1965).
- Martin Fishbein and Bertram H. Raven, "The AB Scales: An Operational Definition of Belief and Attitude," in Fishbein, ed., Readings in Attitude Theory and Measurement (New York: Wiley, 1962).
- O. J. Harvey and R. Ware, "Personality Differences in Dissonance Resolution," Journal of Personality and Social Psychology, 2 (1967), 227-230.
- E. T. Jaynes, "Information Theory and Statistical Mechanics," Psychology Review, 106, 620, (1957).

E. E. Jones and K. E. Davis, "From Acts to Dispositions: The Attribution Process in Person Perception," in L. Berkowitz, ed., Advances in Experimental Social Psychology, Vol. 2 (New York: Academic Press, 1965).

Harold H. Kelley, "Causal Schemata and the Attribution Process," in Edward E. Jones, et.al., eds., Attribution: Perceiving the Causes of Behavior (Morristown, N.J.: General Learning Press, 1972).

Harold H. Kelley, "The Processes of Causal Attribution," American Psychologist, 28 (1973), 107-128.

Michael Mazis and Eugene Klippel, "A Comparison of Four Expectancy-Value Formulations in the Prediction of Consumer Attitudes," unpublished paper, University of Florida, 1973.

Richard Mizerski, "Attribution Theory and Consumer Processing of Unfavorable Information about Products," unpublished doctoral dissertation, University of Florida, 1974a.

Richard Mizerski, "A Test of the Relationship Between Trait and Causal Attribution," in J. Schlinger, ed., Advances in Consumer Behavior, Vol. 2, Proceedings of the Association Consumer Research, 1974b.

James H. Myers and Gregory Warner, "Semantic Properties of Selected Evaluation Adjectives," Journal of Marketing Research, 5 (1968), 409-412.

Darren Newton, "Attribution and the Unit of Perception of Ongoing Behavior," Journal of Personality and Social Psychology, 28 (1973), 28-38.

Thomas S. Robertson and John R. Rossiter, "Children and Commercial Persuasion: An Attribution Theory Analysis," The Journal of Consumer Research, 1 (1974), 13-20.

R. B. Settle, "Attribution Theory and Consumer Behavior," in H. H. Kassarian and Thomas S. Robertson, eds., Perspectives in Consumer Behavior (Glenview, Illinois: Scott, Foresman and Co., 1973).

R. B. Settle and Linda Golden, "Attribution Theory and Advertiser Credibility," Journal of Marketing Research, 9 (1974), 181-185.

C. Shannon and W. Weaver, The Mathematical Theory of Communication (Evanston: University of Illinois Press, 1949).

Susan C. Streufert, "Cognitive Complexity: A Review," Technical Report NO. 2, Office of Naval Research Organization Effectiveness Research Programs, 1972.

# THE IDENTIFICATION OF CONSUMER JUDGMENTAL COMBINATION RULES:

## STATISTICAL PREDICTION VS. STRUCTURED PROTOCOL

C. Whan Park, The University of Kansas  
Charles M. Schaninger, The University of Kansas

### Abstract

Two approaches to the identification of consumer judgmental information processing rules, the statistical prediction method, and the modified structured protocol method were used in a laboratory setting. The degree of consistency between the judgmental rules identified by these two approaches was examined in a number of analyses. A significant degree of consistency between the two methods was found, although the two methods were not perfectly consistent for a substantial portion of the sample. Possible sources of biases in the two methods were examined to evaluate their validity, and suggestions were made for future research.

### Introduction

An increasing number of studies on consumer information processing have been generated since the late 1960's (e.g., Alexis *et al.*, 1968; Bettman, 1970; Haines, 1969, 1970; King, 1969; Russ, 1971; Wright, 1973). These studies can be broadly classified into two types: those concerned with the structural aspects of consumer information processing (Alexis *et al.*, 1968; Bettman, 1970; Haines, 1969; King, 1969), and those concerned with the functional aspects (Russ, 1971; Wright, 1973). The structural approaches are concerned with the structural and sequential aspects of information processing or the paths by which a consumer as a problem solver or decision maker goes about performing the evaluation task, while the functional approaches deal with the specific attribute combination rules used by the consumer for evaluative judgments.

The identification and specification of judgmental rules offers potential not only for increasing the degree of understanding of consumer judgmental processes, but also for the development of practical promotional strategies, providing an appropriate framework to the advertising manager for the adaptation and adjustment of specific advertising plans.

The present study focuses on the functional approach toward consumer information processing. In the next section, several approaches used to identify consumer judgmental models (attribute combination rules) are examined. Two of these approaches are then operationalized to identify the consumer judgmental rules used by subjects to evaluate hypothetical product profiles in an experimental setting.

### Major Approaches Used to Identify the Consumer's Judgmental Rules

A number of approaches have been proposed (Wright, 1973) and partially used in several studies to identify the consumer's judgmental rules. This paper focuses on three major approaches: the unstructured protocol method, the structured protocol method, and the statistical prediction method.

### 1) The Unstructured Protocol Method

Unstructured protocol data are usually obtained by having individuals write down or verbalize their thought processes during decision making. Obtaining unstructured protocol data is a very tedious and time-consuming procedure. Furthermore, once obtained, the analysis is quite informal, i.e., no formal procedure has been developed for inferring and identifying various response models or rules.

For the most part, unstructured protocol data have been analyzed through content analysis. Berelson (1952) has proposed a set of requirements for an effective content coding system: (a) the categories should be mutually exclusive, (b) the categories should be collectively exhaustive, and (c) the categories should record the information as accurately as possible.

Two inherent problems exist in using unstructured protocol data to infer and identify consumer response models (rules). The first is essentially a form of respondent error, stemming from the relative inability of the respondent to recognize and accurately identify his own sequential thought processes. The second is essentially a form of interviewer error, i.e., coding error--that of assigning the respondent's description accurately to appropriate content categories. From the respondent's viewpoint, although the underlying psychological processes which occurred up to the final stages of evaluation may have been quite different, the actual judgmental rules used in the final stages of evaluation (for different underlying processes) are quite similar. This similarity at the final evaluation stage, of different judgmental processes leading up to that point, results in difficulty on the respondent's part in accurately specifying the specific type of judgmental rule(s) actually employed. Thus, the subjects' responses, from the coder's viewpoint, are ambiguous, confounding the two sources of error and possibly increasing the total magnitude of error for this method.

Despite these faults, when no a priori defined models or rules for describing the individual's decision processes have been formulated, the unstructured protocol method is appropriate.

### 2) The Structured Protocol Method

The structured protocol method is a modified version of the unstructured protocol method. It is designed to avoid these problems which arise with the use of the unstructured protocol method. The structured protocol method is constructed as follows:

- 1) Establish the content categories. Each category represents a type of judgmental rule.
- 2) Each content category contains a detailed explanation of a judgmental rule corresponding to that particular category.
- 3) Let the respondent select the content category which represents his thought processes most accurately.

Thus, the respondent's judgmental rule is identified on the basis of his selection of one of several statements (categories) describing judgmental rules, rather than on the researcher's classification.

The structured protocol method is not without inherent problems either. Due to the researcher's specification of "rational" decision making categories, the structured protocol method may measure what the respondent ought to do rather than what he did, i.e., he may choose a more "rational" response because he likes to think himself (or wants to appear) as rational, or because he cannot remember or identify the earlier stages of his underlying cognitive processes.

One way to avoid or reduce the impact of these biases is to combine both the unstructured and structured protocol methods. A pilot study conducted by the author found that response bias problems were reduced by combining both methods: utilizing the unstructured protocol method first, followed by the structured protocol method. This combined form of the unstructured and structured protocol methods was used in this research study, and is referred to hereafter as the Modified Structured Protocol Method. The unstructured protocol method is useful in making the respondent more aware of his own judgmental thinking processes, thus enhancing his ability to accurately select the appropriate statements (categories) representing his thought processes in the structured protocol method which follows. For the modified structured protocol method, the identification of the respondent's judgmental rule is based not on the unstructured protocol data, but on the structured protocol data which follows.

### 3) The Statistical Prediction Method

The statistical prediction method identifies the consumer's judgmental rules by selecting that judgmental rule which yields the highest correlation between the individual's actual and predicted judgment.

Two major interpretational problems may exist with the statistical prediction method: (1) mathematical functional equations developed to represent different underlying cognitive judgmental processes may be algebraic equivalents, and (2) different types of mathematical functions may predict the person's actual judgment with equal effectiveness. These two subproblems are referred to as the "paramorphic" problem by Hoffman (1960). Given the existence of this "paramorphic" problem, it is necessary to institute careful controls on both the informational input (stimulus) and the judgmental evaluation (response) dimensions, when using the statistical prediction method to identify the consumer's underlying judgmental rules.

Four functional equation, corresponding to judgmental models (rules) which are examined by the statistical prediction method in this study, are specified below.

1. Unweighted linear-compensatory model:  $E = \sum_{i=1}^n X_i$
2. Weighted linear-compensatory model:  $E = \sum_{i=1}^n W_i C_i$
3. Conjunctive model:  $E = \min C_i$
4. Disjunctive model:  $E = \max C_i$

where

E = individual's evaluation of an alternative based on its profile of information related to product attributes;

$X_i$  = attribute possession score of the  $i^{\text{th}}$  element in that profile of information;

$C_i$  = attribute possession score of the  $i^{\text{th}}$  salient element (evaluative criterion) in that profile of information;

$W_i$  = subjective importance of the  $i^{\text{th}}$  element;

$\min C_i$  = a minimum of the possession scores on evaluative criteria (salient attributes);

$\max C_i$  = a maximum of the possession scores on evaluative criteria.

Of the approaches used to identify the underlying judgmental rules or models, the modified structured protocol method and the statistical prediction method offer the most promise. Yet, both present possible problems of interpretation and measurement biases. This study utilizes both of these methods, and examines the degree of consistency between them.

Several approaches are used to examine the degree of congruency between the judgmental rules identified by the two methods. The degree of binary consistency between the judgmental rule identified by both methods of identification is examined first, to test the significance of the number of subjects identified as using the same judgmental rule by both the statistical prediction and modified structured protocol methods.

The second approach to examine the degree of consistency focuses on the correlations between predicted and actual judgments. Although the predicted--actual correlation for the judgmental model identified by the structured protocol method cannot be higher than that for the statistical prediction method, it should be higher than the correlations for the other judgmental models not identified by either method. Therefore, the difference between the highest correlation (statistical prediction) and that for the judgmental rule identified by the modified structured protocol method is compared to the differences between the highest relation and the correlations for the other judgmental rules not identified by either method. Thus, the relative predictive accuracies of the two methods are examined.

The third approach used to examine the degree of consistency between the judgmental rules identified by the two methods of identification is an examination of the consistency between the predicted judgments based upon the judgmental rules identified by the structured protocol and statistical prediction methods, using correlational analyses.

The relationships between the judgmental rule identified and the variables of perceived product complexity and prior familiarity are examined for both methods of identification, and the relationships found are compared to each other and to theoretical formulations, to determine the degree to which the relationships found for both methods of identification are consistent with theory and with each other.

### Methodology

Questionnaire data were collected from two hundred and eighty-four junior and senior undergraduate students enrolled in marketing management courses at the University of Illinois and the University of Iowa, in a controlled classroom setting. Each respondent was presented with a questionnaire providing a critically controlled profile of the ratings of eight product



attributes for each of eight hypothetical brands in the product class randomly assigned to him. Based on a previous pilot study, seven consumer products and their attributes were chosen for this study (hamburger, automobiles, toothpaste, stereo cassette decks, exterior trim paint, suntan preparations, and automobile tires). These products were selected to allow a sufficient variation on the dimensions of familiarity and complexity, to be described in the following paragraphs. The respondent was asked to rate the importance of each of the eight attributes for the product class assigned to him, on a seven point scale, ranging from "the most important to me" (7) to "not important to me at all" (1).

Any product criterion (attribute) with an importance score of greater than four was operationally defined as an evaluative criterion. All products with five or more choice criteria were classified as high in complexity; those with less than five as low in complexity. A previous pilot study indicated that this cut-off value was highly associated with respondents' binary classifications of products into high and low complexity groups.

The degree of the individual's prior familiarity with a product was measured by the respondent's agreement with one of three statements assessing the degree to which he had established a set of evaluative criteria for product evaluation. The three statements were the operationalized versions of the three levels of learning phases (extensive, limited and routinized learning) discussed by Howard and Sheth (1969).

Each subject was asked to evaluate along a seven point scale (poor to excellent) for each of eight hypothetical brands in the product class assigned to him. Immediately following each evaluation the subject was asked to specify (in writing) his judgmental processes. Then five structured protocols (the conjunctive, disjunctive, weighted, unweighted linear-compensatory model and none of the above four models) describing each judgmental rule were presented to the respondent. He was asked to identify that particular process which he utilized in evaluating the eight brands.

The critically controlled product attribute profiles were designed to control the informational input dimension on the basis of the following two criteria:

- 1) The rating on each product attribute was designed to reflect no consistent bias (favorable or unfavorable) across all product attributes.
- 2) Not all product attributes which were likely to be important were related favorably (or unfavorably).

These controls were used to minimize the likelihood of different judgmental rules producing the same judgment score, and to make sure that the conjunctive and disjunctive models were as likely to be used as the weighted or unweighted linear-compensatory models.

Judgmental evaluations were obtained for eight brands to control the judgment (response) dimension, in order to achieve a balance between response reliability and statistical reliability. Although it may be argued that a larger number of brands would result in a greater degree of statistical reliability for the correlational analysis used to identify the individual's judgmental model or rule, it was felt that the critical control on the stimulus dimension would reduce this problem. Furthermore, the response reliability resulting from a larger number of brands (judgments) was expected to be substantially lower, due to intentional

refusal to put up with the time-consuming procedure, confused or ambiguous responses resulting from the intensive mental stress required to evaluate an unreasonably large number of brands.

## Results

As stated previously, two different approaches were used to identify the respondent's judgmental rules: the statistical prediction method, and the modified structured protocol method. The criterion used to identify the respondent's judgmental rules for the statistical prediction method was the highest of the four correlations between the actual and predicted judgmental evaluation (of the four models presented previously) on each of the eight brands evaluated. Subjects whose highest correlations were less than .3 were dropped from the analysis presented in this paper because it was felt that none of the four judgmental models adequately represented the judgmental processes of these subjects. These (highest) correlations ranged from .3 to 1.0 and were densely clustered between .7 and 1.0.

TABLE 1  
Cross-tabulation of the judgmental rules identified by the structured protocol and statistical prediction methods

		Judgmental rules identified through statistical prediction				
		CM <sup>5</sup>	ULCM <sup>6</sup>	DM <sup>7</sup>	WLCM <sup>8</sup>	Total
Judgmental rules identified from structured protocol	CM	1 <sup>1</sup>	7	1	9	18
	ULCM	10	37 <sup>2</sup>	3	36	86
	DM	3	5	4 <sup>3</sup>	13	25
	WLCM	9	38	14	60 <sup>4</sup>	121
Total		23	87	22	118	250

<sup>1</sup> expected proportion = .0066, z = -.51

<sup>2</sup> expected proportion = .1197, z = 1.38

<sup>3</sup> expected proportion = .0088, z = 1.22

<sup>4</sup> expected proportion = .2284, z = .435

overall diagonal expected proportion = 0.363584, z = 1.46

<sup>5</sup> CM ---- conjunctive model

<sup>6</sup> ULCM -- unweighted linear-compensatory model

<sup>7</sup> DM ---- disjunctive model

<sup>8</sup> WLCM -- weighted linear-compensatory model

Table 1 presents the cross classification of the judgmental rules identified by the statistical prediction and structured protocol methods. The diagonal frequencies in this table represent the number of respondents demonstrating perfect binary consistency. These observed proportions may be tested against the proportions expected by chance. Two approaches were considered to determine the chance proportions:

- 1) Given the judgmental rule identified by the structured protocol method, the proportion

of .25 may be assigned as that expected by chance for the proportion falling in each diagonal cell.

- 2) The expected proportions may be calculated for each diagonal cell by calculating joint probabilities based upon the product of the proportion observed for a given judgmental rule for the statistical prediction method times that observed for the structured protocol method.

The former approach is reasonable only if the expected overall proportions (column and row subtotal proportions) for the four decision rules are equal. For both methods of identification, the observed proportions were significantly different from .25 at an alpha of .01 for all four decision rules. Therefore, it was decided to use the latter approach to determine the chance proportions, treating the observed proportions as estimates of the chance proportions for each decision rule's identification. These expected proportions for each diagonal cell and for the total diagonal are presented at the bottom of Table 1. None of the z-values was significant for a one-sided alpha of .05, although the z-values for the overall diagonal and for the unweighted linear compensatory model were significant at .10. Thus, only a small degree of consistency between the judgmental rules identified by the two methods was demonstrated for this binary analysis.

The above analysis does not provide insight about the nature of the differences between the judgmental rules identified by the two methods, and is a very stringent examination of the degree of consistency. Even if the rules identified by the two methods were different, the correlations between respondents actual and predicted judgments may be quite similar for the decision rules identified by the two methods, and the predicted judgments for the two methods may also be quite consistent.

If the mean difference score ( $\bar{d}_1$ ) between the correlations based on statistical prediction (the highest correlation) and structured protocol is less than that ( $\bar{d}_2$ ) between the highest correlation and the next highest correlation not identified by either method, and less than that ( $\bar{d}_3$ ) between the highest correlation and the lowest of the correlations not identified by either method, further evidence of the relative consistency between the two methods is obtained.

It is necessary to apply adjusted Fisher's z-transformations to normalize the skewness of the four correlations in order to apply the t-test of differences. The formula used for transformation was  $z = [0.5 \ln((1+r)/(1-r))] - [r/(2(N-1))]$ .

To examine the above hypotheses ( $\bar{d}_1 < \bar{d}_2$ ,  $\bar{d}_1 < \bar{d}_3$ ), t-tests of differences were run for the differences between  $\bar{d}_1$  and  $\bar{d}_2$  and between  $\bar{d}_1$  and  $\bar{d}_3$ . Both of these differences were significant at an alpha of .001 in the expected directions (for  $(\bar{d}_2 - \bar{d}_1)$ ,  $t_{diff} = 3.81$ ; for  $(\bar{d}_3 - \bar{d}_1)$ ,  $t_{diff} = 16.81$ ). Thus, the correlations between predicted and actual judgments for the modified structured protocol method were significantly higher than those for the decision rules not identified by either method of identification, demonstrating consistency with regard to predictive accuracy between the two identification methods.

TABLE 2

Distribution of individual correlations between eight predicted judgments based on the decision rule identified by the statistical prediction method and that identified by the structured protocol method

Correlations	Frequency	Cumulative Percentage
below 0.2	17	6.8
0.2 - 0.29	9	10.0
0.3 - 0.39	19	18.0
0.4 - 0.49	11	22.4
0.5 - 0.59	10	26.4
0.6 - 0.69	17	33.2
0.7 - 0.79	21	41.6
0.8 - 0.89	18	48.8
0.9 - 0.99	26	59.2
1.0	102	100

$$\bar{r}_1 = 0.743$$

$$\bar{r}_2 = 0.566$$

$\bar{r}_1$  is the overall mean correlation for all subjects

$\bar{r}_2$  is the mean correlation excluding those subjects with a perfect correlation ( $r = 1.0$ )

TABLE 3

Crosstabulation between the level of prior familiarity and the decision rule identified by the structured protocol and statistical prediction methods

Judgmental Model-	Familiarity					
	low		moderate		high	
	actual	expected	actual	expected	actual	expected
CM <sup>3</sup>	5 <sup>1</sup>	(4.6)	6	(9.65)	7	(3.74)
	3 <sup>2</sup>	(5.89)	13	(12.33)	7	(4.78)
ULCM <sup>4</sup>	25	(22.02)	49	(46.10)	12	(17.89)
	30	(22.27)	41	(46.63)	16	(18.10)
DM <sup>5</sup>	6	(6.40)	15	(13.40)	4	(5.20)
	7	(5.63)	13	(11.79)	2	(4.58)
WLCM <sup>6</sup>	28	(30.98)	64	(64.86)	29	(25.17)
	24	(30.21)	67	(63.25)	27	(24.54)

<sup>1</sup>frequencies for structured protocol

$$\chi^2_{(prot)} = 8.14, df = 6, N.S.$$

<sup>2</sup>frequencies for statistical prediction

$$\chi^2_{(stat)} = 9.73, df = 6, N.S.$$

$$\chi^2_{(two\ multinomial\ distributions)} = 6.13, df = 11, N.S.$$

<sup>3</sup>CM ---- conjunctive model

<sup>4</sup>ULCM -- unweighted linear-compensatory model

<sup>5</sup>DM ---- disjunctive model

<sup>6</sup>WLCM -- weighted linear-compensatory model

Table 2 presents the distribution of the correlations between the judgments predicted by the decision rule identified by the statistical prediction method and those predicted by that identified by the modified structured protocol method. An examination of these correlations demonstrates that 73.6% of the sample had correlations greater than .621, the critical level of the correlation at the .05 level. Similarly, 55.4% of the reduced sample (excluding the 102 subjects with perfect correlations, the same judgmental model identified by both methods), had correlations greater than the .621 critical level. Thus, there is a relatively high degree of consistency at the individual level between the predicted judgments of both methods of identification.

Cross-tabulation between the level of perceived product complexity and the decision rule identified by the structured protocol and statistical prediction

Judgmental Model	Product Complexity			
	high		low	
	actual	expected	actual	expected
conjunctive model	5 <sup>1</sup> 7 <sup>2</sup>	(6.70) (8.56)	13 16	(11.30) (14.44)
unweighted linear-compensatory model	21 28	(32.00) (32.36)	65 59	(54.01) (54.64)
disjunctive model	10 14	(9.30) (8.18)	15 8	(15.70) (13.82)
weighted linear-compensatory model	57 44	(45.01) (43.90)	64 74	(75.99) (74.10)

<sup>1</sup>frequencies for structured protocol

$$\chi^2_{(prot)} = 12.3, df = 3, P < 0.01$$

<sup>2</sup>frequencies for statistical prediction

$$\chi^2_{(stat)} = 7.9, df = 3, P < 0.05$$

$$\chi^2_{(two multinomial distributions)} = 7.13, df = 7, N.S.$$

Several authors have suggested theoretical formulations concerning the relationships between an individual's information processing modes and his familiarity with the object of concern and the degree of complexity of cognitive structure toward an object (Rosenberg, 1968; Bruner et al., 1962; Ostrom and Brock, 1969). The relationships between judgmental rules with prior familiarity and with perceived product complexity are examined and compared for both methods of identification. Tables 3 and 4 present cross tabulations of the relationships between the judgmental rules identified with prior familiarity and product complexity, respectively, for both the structured protocol and statistical prediction methods. The top value in each row represents the actual frequency for the structured protocol method, while the value directly underneath it represents the frequency for the statistical prediction method; the expected frequencies are presented in parentheses.

Three Chi square values are presented with each table: the value for the judgmental rule identified by the structured protocol method; the value for the statistical prediction method; and a third value representing the significance test of the equality of two Chi square distributions, as suggested by Mood et al. (1974). This modified Chi square test is a form of the test for the equality of two multinomial distributions, and can be expressed as follows:

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^3 \sum_{k=1}^4 [N_{ijk} - ((N_{1jk} - N_{2jk})/2)]^2 / [(N_{1jk} + N_{2jk})/2]$$

with (jk) - 1 df

$N_{ijk}$  = the observed frequency in cell jk for distribution i;  
 i = the number of distributions compared  
 j = the number of columns (levels of familiarity)  
 k = the number of rows (types of judgmental rules)

If the relationships between the judgmental rule used by the individual with prior familiarity and product complexity are equivalent (consistent) for both methods of identification, further evidence is provided for the relationships (between judgmental models with familiarity and complexity), and for the consistency between the two identification methods, although aggregate analysis does not necessarily represent the underlying relationships at an individual level analysis.

Theoretical speculations (Rosenberg, 1968; Park, 1974) suggested that the unweighted linear-compensatory model was more likely to be used when an individual is unfamiliar with a product (class), and that the weighted linear-compensatory model was more likely to be used when an individual is quite familiar with a product (class). Although neither Chi square value was significant, the results of both methods were in the direction of the theoretical formulations. The modified Chi square value testing the equality of the two distribution was not significantly different. A closer examination of the two results reveals that the statistical prediction method exhibited stronger results for those classified as low in familiarity (they were more likely to use the unweighted and less likely to use the weighted linear-compensatory model), while the structured protocol method exhibited stronger results for those high in familiarity (they were more likely to use the weighted and less likely to use the unweighted linear-compensatory model).

Bruner et al. (1962) introduced two types of information processing, careful-but-slow strategy and risky-but-fast strategy, when referring to the conjunctive and the disjunctive models. It is strongly believed that the risky-but-fast strategy is not an appropriate strategy for the person whose cognitive structure toward a product is relatively complex (when a large number of product attributes are viewed as important). An examination of the relationships for both methods of identifying decision models reveals that significant relationships were found for both methods. But, only the results for the statistical prediction method were in accord with the theoretical speculations. The Chi square value for the equality of the two distributions was not significant, although the interpretation of the findings of the two Chi square distributions would be different. The data for the structured protocol method did not indicate a relationship between complexity and use of the disjunctive model, while the data for the statistical prediction model did. The

results of the structured protocol model indicated that those high in complexity were more likely to use the unweighted and less likely to use the weighted linear-compensatory model than those low in complexity. A comparable tendency was found with regard to use of the unweighted linear-compensatory model for the statistical prediction method, but not for the weighted model.

It should be noted that subjects were classified as high in complexity if five or more (out of eight) attributes were rated as high (above 4 on a 7 point scale) in importance. One feasible explanation for the differences in use of the unweighted and weighted linear-compensatory models between those grouped as high versus low in complexity may be that those high in complexity dealt with a larger number of "evaluative" criteria which they perceived as being more equal in importance (all high in importance) than those classified as low in complexity. The greater number of attributes considered as important, and the lesser degree to which the importance ratings differed across attributes, may have led to the greater reported usage of the unweighted and lower reported usage of the weighted linear-compensatory models by the subjects for the structured protocol method. The lack of difference for the weighted model for the statistical prediction method may have been due to the statistical artifact problem.

#### Discussion

Weak evidence of consistency between the results of the statistical prediction and structured protocol methods was found when their binary consistency was examined. However, when the correlations between predicted and actual judgments were examined, it was found that the correlations for the decision rules identified by the modified structured protocol method were higher than those for the other decision rules not identified by either method. It was also found that high positive correlations existed between the predicted judgments for the judgmental rules identified by the two methods. Both of these findings indicate a significant degree of consistency between the decision rules identified by the modified structured protocol and statistical prediction methods. No significant differences were found between the relationships of judgmental rules with the theoretical constructs of prior familiarity and perceived complexity for the two methods of identification, although examination of the two sets of relationships could lead to different conclusions about the nature of the relationships. The results for the statistical prediction method were more consistent with theoretical speculations.

More detailed information concerning the underlying judgmental processes is needed to determine which of the two methods of identification is more valid. Both approaches have sources of possible bias. For the modified structured protocol method, respondents may be unable to differentiate one model from others while tracing back their thought processes; or their preferences for "rational" or thought oriented decision-making rules may influence their responses; or the set of judgmental rules used may be less than comprehensive. The statistical prediction method may be subject to biases due to: the paramorphic and statistical artifact problems (Hoffman, 1960); incomplete sets of evaluative criteria presented by the researcher; or halo effects of researcher-defined "salient" criteria. Incomplete specification of the full set of judgmental rules, or combinations of rules, as suggested by a number of authors (Russ, 1971; Wright, 1975; Dawes, 1964; etc.) may lead to biases in the results of both

the statistical prediction and modified structured protocol methods.

A more process oriented approach may overcome some of these biases. Modifications of the discrimination net approach (Bettman, 1970; Clarkson, 1961; Swinth, 1975) may prove fruitful if modified to allow the examination of the specific impact of each stimulus on evaluations or choices and, if modified, to allow the incorporation of configural processes or compensatory models. The determination of the salience of various stimuli would be more accurate when examined in a decision net process oriented approach than with questionnaire measures such as importance scales. The response bias problems due to preference for "rational" judgmental rules and to difficulty in differentiating one model from others, inherent in the structured protocol method may also be eliminated with such a process oriented approach. Such an approach would also provide information about the sequential or multi-stage aspects of decision processes which utilize combined characteristics of two or more models. Although such process-based information is costlier and more time consuming to collect, and harder to make generalizations with, such information would permit the evaluation of the validity of other identification approaches such as the statistical prediction and structured protocol methods.

Wright's (1975) approach of training subjects to use different attribute combination rules may also be quite fruitful if combined with actual choice situations where respondents are asked which combination rule, or sequential combination of rules they actually used. This approach allows the researcher to control for the number of alternatives, and in actual choice situations, allows the individual to specify which evaluative criteria (and their number) he actually used. Information obtained in such a setting would alleviate some of the biases due to measurement problems, encountered with the statistical prediction method in this study.

Furthermore, after such training, it would be possible to utilize questionnaire measures to determine the number of alternatives, number of salient attributes, and sequential combination rules reported as typically used, by the individual for a number of product classes, and to investigate their interrelationships.

#### References

- M. Alexis, G. Haines, and L. Simon, "Consumer Information Processing: The Case of Women's Clothing," Proceedings (American Marketing Association Fall Conference, 1968), 197-205.
- B. Berelson, Content Analysis in Communication Research (Glenco, Illinois: The Free Press, 1962).
- J. R. Bettman, "Information Processing Models of Consumer Behavior," Journal of Marketing Research, 7 (August, 1970), 370-376.
- J. S. Bruner, J. J. Goodnow, and G. A. Austin, A Study of Thinking (Science Editions, 1962).
- G. P. Clarkson, "Decision-Making in Small Groups: A Simulation Study," Behavioral Science, 13 (July, 1968), 288-305.
- G. H. Coombs, Theory of Data (New York: John Wiley and Sons, 1964).

- R. M. Dawes, Toward a General Framework for Evaluation (Ann Arbor: University of Michigan, Department of Psychology, 1964).
- R. M. Dawes, "Social Selection Based on Multidimensional Criteria," Journal of Abnormal and Social Psychology, 68 (January, 1964), 104-109.
- L. R. Goldberg, "Simple Model or Simple Process? Some Research on Clinical Judgments," American Psychologist, 23 (July, 1968), 483-496.
- G. H. Haines, Jr., "Process Models of Consumer Decision Making," presented at the Third Annual Conference on Buyer Behavior, Columbia University, May 22 and 23, 1969.
- P. J. Hoffman, "The Paramorphic Representation of Clinical Judgment," Psychological Bulletin, 57 (March, 1960), 116-131.
- J. A. Howard and J. A. Sheth, The Theory of Buyer Behavior (New York: John Wiley and Sons, 1969).
- R. H. King, "A Study of the Problem of Building a Model to Simulate the Cognitive Processes of a Shopper in a Supermarket," in G. H. Haines, Consumer Behavior: Learning Models of Purchasing (New York: Free Press, 1969), 22-67.
- A. Mood, F. A. Graybill, and D. C. Boes, Introduction to the Theory of Statistics, Third Edition (New York: McGraw-Hill Publishing Company, 1974).
- T. Ostrom and T. Brock, "Cognitive Bonding to Central Values and Resistance to a Communication Advocating Changes in Policy Orientation," Journal of Experimental Research in Personality, 1 (July, 1969), 30-41.
- C. Whan Park, "An Exploration of the Consumer's Judgmental Rules," unpublished Doctoral Dissertation, University of Illinois, 1974.
- M. J. Rosenberg, "Discussion: Information Processing and the Evaluation of New and Old Subjects," in R. P. Abelson et al., Theories of Cognitive Consistency: A Sourcebook (Chicago: Rand McNally, 1968).
- F. A. Russ, "Evaluation Process Models and the Prediction of Preference," Proceedings (Second Annual Conference, Association for Consumer Research, 1971), 256-261.
- R. L. Swinth, "A Decision Process Model for Predicting Job Preferences," Faculty Working Paper, School of Business, University of Kansas, 1975.
- P. L. Wright, "The Simplifying Consumer: Perspectives on Information Processing Strategies," presented at the American Marketing Association, Doctoral Consortium, East Lansing, Michigan, August, 1973.
- P. L. Wright, "Analyzing Consumer Judgment Strategies: Paradigms, Pressures and Priorities," Faculty Working Paper #94, College of Commerce and Business Administration, University of Illinois at Urbana-Champaign, 1973.
- P. L. Wright, "Consumer Choice Strategies: Simplifying vs. Optimizing," Journal of Marketing Research, 12 (February, 1975), 60-77.

WORKING WIFE VS. NON-WORKING WIFE FAMILIES:

A BASIS FOR SEGMENTING GROCERY MARKETS?

Susan P. Douglas<sup>1</sup>

Centre d'Enseignement Superieur des Affaires  
Jouy-en-Josas, France

Abstract

The growth in the number of working wife families is widely acknowledged to be one of the most important social trends of the era. Yet little attention has been paid to examining its implications for shopping behavior patterns and marketing strategies. The study outlined in this paper was intended to explore this issue, and in particular whether the wife's employment status is likely to provide a basis for segmenting markets for grocery products.

Introduction

Over the past half century the proportion of working wife families has increased by leaps and bounds. Whereas in 1910, only one out of five wives worked, today the proportion is close to three out of five (Ferris, 1971). The consequences for family behavior and life-style patterns once the wife's main working hours are removed from the home are overwhelming in magnitude (Blood and Hamblin 1958, Hedges and Barnett 1972, Heer 1958, Linden 1973, Nye and Hoffman 1963). As the Presidential Commission on Manpower in 1972 suggested, "perhaps no other single change in family life has affected so many families in as relatively short a time as has the movement of married women into the labor force" (U.S. Department of Labor 1972). Despite the importance of this social trend, scant attention appears to have been paid to examining its implications for marketing strategies. Relatively little interest seems, for example, to have been shown in assessing whether working wife families have specific needs and interests which provide opportunities for developing new products and services, or for emphasizing certain benefits and appeals.

There are two major reasons for thinking that this neglect of the working wife families is unwarranted. In the first place, time constraints on the performance of household duties such as shopping, cooking, cleaning, and in particular on the time at which these tasks are performed in working wife families, suggest that different strategies are likely to be developed to cope with these problems. One hypothesis is that working wives will show greater interest than non-working wives in time-saving products and services such as take-out dinners, laundry services, and instant dusting spray (Management Review 1967).

---

<sup>1</sup> Financial support for this project was provided by the Centre d'Enseignement Superieur des Affaires, Jouy-en-Josas, and the Marketing Science Institute, Cambridge, Mass. The author wishes to thank Professor Frank Carmone (University of Hong Kong), Professor Yoram Wind (University of Pennsylvania), and Professor Bernard Dubois (C.E.S.A., Jouy-en-Josas) for their assistance in this project, and Ms Jane Ross (Marketing Science Institute) for performing the computer runs associated with the study.

Secondly, differences in attitudes among working and non-working wives towards female roles are likely to influence their behavior. The dual roles of the wife in her employment and in the home, imply that the degree of involvement in various "homemaker" roles such as "Chief Cook", "Mother", "Housekeeper", will differ from the non-working wife for whom such roles are the focal point of her existence (Hartley 1960). This does not necessarily imply rejection of homemaker roles by the working wife, but rather that she seeks other opportunities for self-fulfillment in activities outside the home (Bailyn 1970, Powell 1960). Consequently, the time and effort devoted to homemaker roles, as well as concepts of appropriate behavior in these roles, may differ from those of the non-working wife. Stimulation received from involvement in a job may also generate greater interest in new and different products, and emphasis on different product benefits; for example, more concern with originality, or less concern with the preference of family members. Nonetheless, working wives, and equally non-working wives, may not all have similar attitudes toward homemaking. Motivations for working or for not working differ (Hoffman 1963b). Some women work primarily from financial necessity and remain predominantly involved in homemaker roles. Equally some non-working women solve frustrations with the domestic routine by involvement in non-paid social and charitable activities rather than paid employment. Consequently, they are open to an equally wide sphere of contacts and relationships as working women, and may have as little time or energy to devote to household duties.

Scant attention has been paid to examining the impact of the wife's employment status on purchase behavior. Some comparisons have been made of purchase behavior in working and non-working wife families. These have found, for example, that working wives tend to be more efficient in organizing household tasks; they make fewer shopping trips, and are more likely to be accompanied by their husbands than non-working wives (Anderson 1972, Baldwin and Lunn 1972, Hoffman 1963a, Lunn, Baldwin and Dickens 1972). On the other hand, there appears to be little evidence to indicate that working wives are more interested in convenience products than non-working wives (Anderson 1972).

Such studies barely scratch the surface of the working wife phenomenon. Typically, they compare working and non-working wife families without making a distinction between part-time and full-time working wives, nor taking into account the interaction between employment status and other factors such as family income, age, number of young children. Nonetheless, these are likely to have an important bearing on the wife's time availability and interest in convenience products. If, for example, non-working wives have several young children at home, or are heavily involved in activities outside the home, they are likely to have equivalent or greater time pressures to part-time working wives; equally, the older non-working wives may have greater

discretionary income to spend on convenience products and services than young working wives. In addition, little research has been undertaken to examine differences in attitudes among working and non-working wives, and the extent to which these influence their purchase behavior.

The purpose of the exploratory study described in this paper was to investigate the effect of the wife's employment status on grocery purchasing behavior, taking into consideration differences in other socio-economic and demographic characteristics among working and non-working wives and in attitudes toward various women's roles. The two major research questions were as follows:

1. Do working wives appear to have substantially different shopping behavior patterns from non-working wives likely to provide an adequate basis for segmenting grocery markets? Or are observed differences merely a reflection of differences with regard to other factors such as education, income, or the number of children at home?
2. Are there important attitudinal differences within the two groups which lead to different purchase behavior patterns, and suggest that a finer level of analysis based on sub-segments within each group is likely to be more relevant for management purposes?

First the data base and the research design for the study are briefly presented, followed by a discussion of the research findings relating to the two key issues. Some issues in the investigation of the impact of the wife's employment status on purchase behavior are then examined and guidelines for further research proposed.

#### Research Methodology

##### Data Base

In the study a questionnaire was administered by personal interview to 106 married women in the Philadelphia and Boston areas. All respondents had been married at least two years and were between the ages of 25 and 50. Approximately half the sample (49) was not gainfully employed outside the home, while the other half (57) was composed of women who had been employed full-time for at least eighteen months. Their jobs included a variety of occupations such as lawyers, doctors, researchers, secretaries, waitresses and sales clerks. The sample was selected by a quota procedure and both the working and non-working groups were stratified by age, income and number of children. The objective was thus to obtain two groups "matched" in terms of background characteristics, rather than two groups which were representative of the working and non-working wife populations.

The structure of the sample and its small size clearly limits the extent to which any substantive conclusions can be drawn concerning differences in purchase behavior of working and non-working wife families, particularly insofar as differences within the two groups are concerned. The purpose of the study was not, however, to compare these differences, but rather to isolate the impact of the wife's employment status from that of other related variables, such as income, stage in family life-cycle, number of young children at home. Consequently, given research budget constraints, the alternative of small matched samples controlling for these variables was preferred to that of larger but less comparable samples.

In the questionnaire, respondents were asked about grocery purchase behavior; attitudes to female roles; and background characteristics of the family. The grocery purchase questions examined 1) interest in various product benefits when purchasing grocery items; 2) frequency of purchasing selected convenience products and services; and 3) the organization of shopping activities, for example, frequency of shopping in different types of stores and husband participation in shopping activities. Attitudes toward four aspects of a woman's role were examined: a) home and family orientation, b) cooking, shopping and other household tasks, c) self-perceptions and concepts, and d) social interaction. The background questions included a number describing the home, such as the type of household appliances owned, as well as the standard family background questions relating to income, age, and education of husband and wife, etc.

##### Data Analysis

Following the two major research questions the analysis was divided into two parts (Figure 1). First, overall differences in grocery purchase behavior between the two groups were examined. Then, differences in attitudes towards female roles within each group were investigated, together with the purchasing behavior patterns of different attitudinal sub-groups.

Overall differences between working and non-working women. After a preliminary series of Chi square tests of all three sets of grocery purchase variables, two, use of convenience products and grocery shopping, were submitted to a multivariate analysis of covariance (BMD 12V, Green and Tull 1975) to test whether apparent differences between working and non-working wives were significant after adjustment for differences in 11 socio-economic characteristics of the two groups (the covariates).

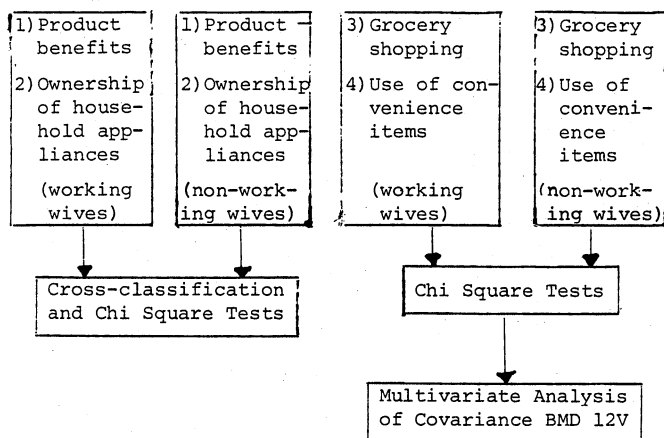
Within group analysis. In the second stage of the analysis, sixteen key attitudinal variables (four from each of the four sets of attitudes) were identified by means of a series of factor analyses<sup>2</sup>(Table 1). Cluster analyses of these variables using the Howard Harris clustering program (Howard and Harris 1966) were then conducted separately for the working and non-working wife groups; two major sub-groups were identified among working wives and three among non-working wives. Differences in grocery purchase behavior of these groups were then examined by means of discriminant analyses.

<sup>2</sup>First a separate factor analysis for each of the four sets of attitudinal characteristics was conducted (BMD 08M). The variable with the highest loading on the first four factors in each factor analysis was then retained for the cluster analysis. The total number of factors identified precluded the use of all the factors, but the first four in each case accounted for between 60.8 - 80.2% of the total variance. For further discussion of the rationale for this procedure see Susan Douglas and Patrick Le Maire (1974). The factor analyses were conducted jointly for working and non-working groups after preliminary analysis based on separate factor analyses resulted in the identification of highly similar underlying structures suggesting that the joint analysis could appropriately be used to summarize the data for both groups.

FIGURE 1

Flow Chart of Analysis

a) primary group analysis



b) within group analysis

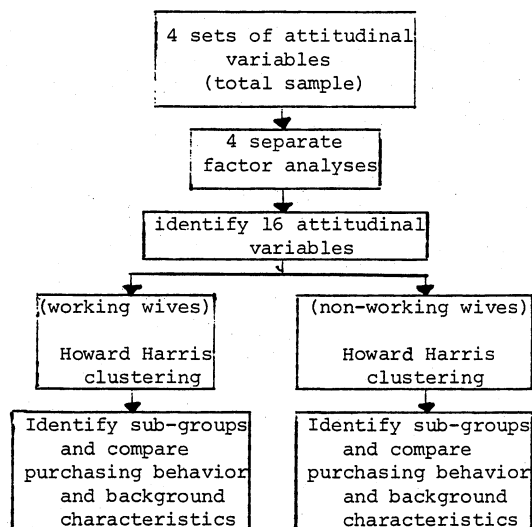


TABLE 1

The 16 Key Attitudinal Statements

<p><u>Home and Family Orientation</u></p> <p>*A woman's place is in the home</p> <p>*I am uncomfortable when my house is not completely clean</p> <p>*It is right that women should not be treated equally to men</p> <p>I enjoy most forms of housework</p> <p><u>Shopping and Cooking</u></p> <p>I like to shop in a store where I feel at home</p> <p>I am more interested in food products than most people</p> <p>I love to bake cakes and frequently do so</p> <p>*I am generally the first of my friends to buy a new product</p> <p>*Variables discriminating between working and non-working wives in a stepwise discriminant analysis.</p>	<p><u>Self-Perceptions and Concepts</u></p> <p>*I often feel I spend too much on clothes</p> <p>*I invariably buy the latest fashion</p> <p>I always feel rushed and in a hurry</p> <p>I believe in salvation, that is eternal life</p> <p><u>Social Interaction</u></p> <p>I like to feel attractive to men</p> <p>I think every woman should participate in church and neighborhood activities</p> <p>I spend a lot of time reading</p> <p>I like going to parties</p>
--	--

Results

Differences Between Working and Non-Working Wives

Some differences emerged in grocery purchasing behavior of the working and the non-working wives but these were not always along the lines hypothesized.

Grocery shopping patterns. As Table 2 indicates there are some slight differences between the two groups in the way shopping activities are organized; particularly insofar as the working wives go less frequently to neighborhood stores, (possibly implying fewer mid-week trips), and tend to make somewhat more use of husbands. These differences were substantially more

marked in the unadjusted means, and tend to be affected by the wife's age and the socio-economic status of the family.

Use of convenience products and services. Here the results confirm those of other studies insofar as the working wives do not have any systematic tendency to use time-saving products and services more frequently than the non-working wives. In fact the latter appear to be heavier users of take-out dinners and several other products such as instant dusting spray and baked goods.



Reasons for buying grocery products. The working and non-working wives also appear to consider similar factors important when buying grocery products for regular meals. The only significant difference occurs in

relation to "something different" which, contrary to expectations, is more important to the non-working than working wives (Table 3). Reasons for buying when entertaining company are also similar, but are not reported here.

TABLE 2

Results of The Multivariate Analyses of Covariance For Grocery Shopping and Use of Convenience Products and Services (adjusted means and approximate F statistics for the multivariate analysis)<sup>3</sup>

	Working Wives	Non-Working Wives
Grocery shopping variables		
All shopping trips	4.3	4.3
Shops in large supermarket	3.9	3.9
Shops in neighborhood supermarket	1.9	2.6*
Shops in corner grocery store	2.4	2.5
Shops in farmer's market	1.7	1.8
Husband accompanies on grocery shopping trips	1.9	1.9
Husband does major shopping trips	1.7	1.5
Husband does occasional shopping trips	2.2	2.2
F statistic for the multivariate analysis:		1.2

	Working Wives	Non-Working Wives
Use of convenience products and services		
Takeout dinners (pizza, Chinese food)	2.3	2.6*
Instant desserts (not ice cream)	1.9	1.9
Canned main dishes	1.4	1.3
Baked goods (pies and cakes bought from the store)	2.8	2.9
Frozen main dishes, TV dinners	2.1	1.9
Cold cuts (ham, salami, roast beef)	3.3	3.4
Laundry services (sheets)	1.4	1.2*
Paper plates and cups	2.6	2.4
Aerosol carpet cleaner	1.3	1.4
Instant dusting spray (Pledge)	2.7	2.8
F statistic for the multivariate analysis:		0.6

\*Univariate statistic significant at the .05 level.

TABLE 3

A Comparison of Reasons for Selecting Grocery Products and of Ownership of Household Appliances for Working and Non-Working Wives

	Working Wives	Non-Working Wives
Reasons for buying grocery products	% selecting for regular family meal	
Low price	41.1	39.6
Liked by husband	59.6	67.3
Liked by children	78.9	87.9
Economical	64.3	61.3
Good quality	93.0	85.7
Easy to prepare	46.6	44.9
Always tastes good	63.0	49.0
Recommended by friends	9.1	20.4
No leftovers	7.3	10.4
Something quick	44.6	42.9
Something different	37.5	61.7**

\*\*Chi square statistic significant at .05 level.

	Working Wives	Non-Working Wives
Ownership of household appliances	% Owning	
Dishwasher	61.4	71.4
Clothes washer	84.2	98.0**
Clothes dryer	75.4	91.8**
Self-cleaning oven	17.5	16.3
Self-defrosting refrigerator	59.6	79.6**
Electric carving knife	40.4	63.3
Blender	91.2	93.9
Mixer	94.7	93.9
Separate freezer	17.5	36.7

<sup>3</sup> A five point scale was used to rate frequency of shopping and usage of convenience products; the higher the mean, the more frequent the activity or usage.

Household equipment and domestic help. There is also little indication that the working wives rely more on the use of domestic help or on mechanical labor saving devices to aid in their cooking and housekeeping chores. On the contrary, the non-working wives appear to be better equipped than their working counterparts, especially with washing machines and dryers, freezers, electric carving knives and self-defrosting refrigerators (Table 3). Perhaps, since non-working women spend more time in the home, they attach more importance to possessing these appliances than do working women.

Although the small sample size precludes any definitive conclusions, the wife's employment status appears to affect the way in which shopping activities are organized but not the time devoted to meal preparation and other household tasks. Time pressure problems appear to be solved either by a sharing of tasks with other family members or simply by the wife's absorption of the increased work load rather than by recourse to outside sources of help and shifting of the burden to the commercial sector.

Consequently, families in which the wife is employed are unlikely to be prime targets for time-saving convenience products and services or pioneers in the use of such products. Wives who do not work appear to be equally interested in such items, possibly because they too are pressured for time due to involvement in social or neighborhood activities. On the other hand, since the working wives appear to seek solutions for saving time spent shopping, emphasis by retailers on speed of service and home delivery services could prove an attractive selling point for them.

The increased participation of husbands in working wife families in grocery shopping and particularly in assuming the major responsibility for these tasks, also suggests that the relevant decision-making unit and hence appropriate promotional target is different when the wife is employed. While in non-working wife families, focus predominantly on the wife is likely to be appropriate, in working wife families inclusion of both husband and wife appears desirable. This does, however, depend on the extent to which the husband influences product and brand decisions. Examination of this aspect is therefore needed before any definitive conclusions can be reached.

#### Differences Among Working and Non-Working Wife Sub-Groups

In the second phase of the analysis attention was centered on examining differences in attitudes towards various female roles and related purchase behavior patterns within the two groups. Again the small size of the groups identified implies that findings should be viewed as exploratory, and that considerable further investigation and validation of apparent differences is required.

Working wives. Among the working wives two highly distinct groups were identified; their profiles are shown in Table 4. The "progressive egalitarian" group was characterized predominantly by their liberated attitude toward a woman's role. Although not houseproud, they enjoyed cooking, but were less innovative and fashion-conscious than the other group. The second group of "fashionable traditionalists" had traditional attitudes toward female roles, and were houseproud. While interested in fashion they tended to feel somewhat guilty about the amount they spent on clothes.

TABLE 4

#### Key Discriminating Attitudinal Variables for The Working Women Sub-Groups

(means on 6 point Likert scale)

Variable	Progressive Egalitarians (33)	Fashionable Traditionalists (24)
It is right that women should not be treated equally to men	1.0	2.8
I invariably buy the latest fashion	2.2	4.8
A woman's place is in the home	1.2	2.1
I am generally the first of my friends to buy a new product	2.1	3.4
I often feel I spend too much on clothes	1.6	2.8
I love to bake cakes and frequently do so	4.5	3.4
I like to shop in a store where I feel at home	5.1	4.8
I am uncomfortable when my house is not completely clean	2.7	4.8
I believe in salvation, that is, eternal life	2.6	3.5

A comparison of the grocery purchase behavior of these two groups revealed some differences (Table 5), though not necessarily what one might expect from their attitudes. The "liberated" views of the "egalitarians" suggest that they may attach less importance to being a "good" housewife than other working wives and hence devote less effort to housekeeping activities. On the contrary, however, they appear to make significantly less use of all types of convenience products and services, and to shop more frequently in the time-consuming farmers' markets than the "traditionalists". Perhaps this results from their greater concern with economy.

Some partial explanation for these patterns can be found in the socio-economic characteristics of the two groups. Members of the "progressive" group are generally better educated and have higher status jobs than the "traditionalist" group (Table 6). Consequently they may have higher achievement levels and attach greater importance to performing effectively in both domestic and career roles. The "traditionalists", on the other hand, may work predominantly for financial reasons, and feel little need to compensate for the time spent away from home by devoting effort to meal preparation and other household chores.

Non-working wives. The profiles of the three attitudinal sub-groups identified among non-working wives were not as distinctive as those of the working wife groups, but again differed primarily in attitudes to a woman's role and interest in clothes (Table 7). One group - "the disorganized homemaker" - was characterized primarily by her tendency always to feel rushed. She also had somewhat conservative role perceptions

and showed little interest in clothes or food products. The "liberated socializer", on the other hand, rejected the traditional view of a woman's role and enjoyed reading and going to parties but tended to have guilt feelings about her expenditures on clothes. The "conservative homebody" had highly traditional role perceptions but was outgoing and interested in fashion and new food products.

Examination of the grocery shopping behavior of the three groups showed some differences, particularly in the frequency of shopping, and husband participation

in shopping, but the patterns were not as clearcut as for the working wife groups (Table 8).

Although further investigation and validation of both the working and non-working sub-groups is clearly necessary, attitudes toward a woman's role emerge as an important factor differentiating among women in both groups. Attention might therefore usefully be focused on further investigating the role of such factors and their relationship to interest in different types of products as well as other aspects of grocery shopping behavior among working and non-working wives.

TABLE 5

Key Discriminating Grocery Shopping Variables  
For Two Working Women Attitudinal Groups<sup>4</sup>

<u>Reasons for buying grocery products for regular family meal</u>	<u>Progressive</u>	<u>Fashionable</u>
	<u>Egalitarians</u>	<u>Traditionalists</u>
	<u>% selecting characteristic</u>	
Low Price	54	22*
Economical	76	48*
Always tastes good	49	84*
*Chi square statistic significant at the .05 level		
<u>Use of convenience products and services</u> (means of 5 pt. scale)		
Takeout dinners (Pizza, Chinese food)	1.9	2.6**
Instant desserts (not ice cream)	1.3	2.6**
Baked goods (bought from the store)	1.9	3.6**
Frozen main dishes, TV dinners	1.7	2.4**
Average for all products	<u>1.7</u>	<u>2.4</u>
<u>Grocery shopping variables</u> (means on 5 pt. scale)		
All shopping trips	4.2	4.3
Shops in large supermarket	3.8	4.0
Shops in neighborhood supermarket	1.9	2.0
Shops in corner grocery store	2.5	2.5
Shops in farmer's market	2.2	1.1**
Husband accompanies on grocery shopping trips	2.1	1.9
Husband does major shopping trips	1.8	1.5**
Husband does occasional shopping trips	2.2	2.0
**statistically significant in the discriminant analysis		

<sup>4</sup>Variables for which the means are similar are omitted.

TABLE 6

Differences in Occupational Status of the Two Working Wife Groups

	<u>Husband's Occupation</u>		<u>Wife's Occupation</u>	
	<u>Progressive</u> <u>Egalitarians</u>	<u>Fashionable</u> <u>Traditionalists</u>	<u>Progressive</u> <u>Egalitarians</u>	<u>Fashionable</u> <u>Traditionalists</u>
Upper	48.5%	12.5%	18.2%	8.3%
Middle	45.5%	45.8%	72.7%	45.8%
Lower	<u>6.1%</u>	<u>41.7%</u>	<u>9.1%</u>	<u>45.8%</u>
	100.0%	100.0%	100.0%	100.0%

TABLE 7

Key Discriminating Attitudinal Variables for  
The Three Non-Working Wife Groups  
(means on 6 pt. Likert scale)

<u>Variables</u>	<u>Disorganized Homemaker</u> (16)	<u>Liberated Socializer</u> (15)	<u>Conservative Homebody</u> (18)
A woman's place is in the home	2.3	1.1	4.6
I often feel I spend too much on clothes	1.5	3.8	2.1
I always feel rushed and in a hurry	5.0	3.0	3.3
It is right that women should not be treated equally to men	2.2	1.2	1.2
I am more interested in food products than most people	1.8	2.0	3.5
I spend a lot of time reading	4.6	5.2	4.0
I like going to parties	4.1	5.2	5.5
I invariably buy the latest fashion	3.1	3.3	3.6

TABLE 8

Key Discriminating Grocery Purchase Variables for  
The Three Non-Working Groups<sup>5</sup>

	<u>Disorganized Homemaker</u>	<u>Liberated Socializer</u>	<u>Conservative Homebody</u>
<u>Reasons for buying grocery products</u>	<u>% selecting characteristic</u>		
Low price	40	33	40*
Liked by husband	75	66	61*
Easy to prepare	68	20	44*
Always tastes good	56	26	61*
*Chi square statistic significant at the 0.5 level.			
<u>Use of convenience products &amp; services</u> (means on 5 pt. scale)			
Instant desserts (not ice cream)	1.3	2.3	2.1*
Frozen main dishes, TV dinners	1.6	1.7	2.5*
Instant dusting spray (Pledge)	<u>2.7</u>	<u>2.3</u>	<u>2.3*</u>
Average for all products	1.9	2.0	2.3
<u>Grocery shopping variables</u> (means on 5 pt. scale)			
All shopping trips	4.4	4.2	4.0*
Shops in large supermarket	4.1	3.5	4.2*
Shops in neighborhood supermarket	2.7	2.3	2.8
Shops in corner grocery store	1.8	2.8	2.3*
Shops in farmer's market	1.5	2.2	1.4
Husband accompanies on grocery trips	1.5	2.1	2.0
Husband does major shopping trips	1.3	1.6	1.6
Husband does occasional shopping trip	2.1	2.6	1.9*
*statistically significant variables in stepwise discriminant analyses			

<sup>5</sup>Variables for which means are similar are omitted.

## Conclusions

Although at this stage it would clearly be premature to draw any substantive conclusions, the findings suggest a number of guidelines for future research.

In the first place, the results to a large extent confirm those of previous studies which suggest that working wives do not differ significantly from non-working wives in the use of convenience products and services, nor in interest in different product benefits when buying grocery products. This is particularly interesting insofar as the hours worked by the wife, her age and number of young children were controlled for, and hence the lack of differences between working and non-working wife families cannot be attributed to such factors. On the other hand, certain differences emerge in the organization of shopping activities in working and non-working wife families which could usefully be further investigated, particularly in relation to the impact of the wife's employment on husband involvement in shopping activities.

Secondly, focus on different sub-groups among working and non-working wives appears to be a promising line for future research. In addition to examining attitudinal differences, the impact of factors such as occupation and education level on purchase behavior needs to be considered particularly in the case of working wives. In this respect investigation of differences in the purchasing behavior of women in different status jobs, and with different work motivations, for example, between professional "career" women and women with lower status jobs merits attention.

A third issue concerns the relationship between attitudes toward female roles and purchase behavior. While groups with different attitudes appear to be characterized by different purchase behavior, the nature of the relationship between attitudes and behavior is not clear. Intuitively, one would expect traditional female role perceptions to lead to a rejection of convenience products and services, yet the reverse relationship occurs both among working and among non-working wives. Possibly home-oriented women have lower levels of self-achievement, and are more ready to accept the easy solution. Further exploration of this issue is also required.

In many respects the study raises more questions than it solves. Yet it helps to pinpoint some key issues, and to demonstrate the complexity of the working wife phenomenon and its implications for shopping behavior. The fact that a woman is involved in employment outside the home does not necessarily imply that she will devote less effort to homemaking activities than her non-working counterpart; conversely, a non-working wife is not necessarily highly committed to her domestic role. Nonetheless involvement in a full-time job imposes time constraints on the performance of household duties which have consequences for purchase behavior. In addition although involvement in a job is associated with different attitudes to homemaker roles, these vary among different categories of working wives; and the precise impact on purchase behavior remains unclear.

Further clarification of these issues should be a major concern at this time. As the women's liberation movement gains more widely based support, the invasion by women of upper status jobs is likely to continue. In addition, the shift of married women into the labor force shows no signs of abating. Mastering the implications of these major social forces is thus likely to be critical to understanding future developments in grocery shopping patterns.

## References

- Beverlee Anderson, "Working vs. Non-Working Women: A Comparison of Shopping Behavior", Proceedings of the A.M.A. Fall Conference (Chicago: A.M.A. 1972).
- E. S. Baldwin and J. A. Lunn, "Researching the Changing Consumer", Proceedings of the ESOMAR Congress (Cannes: ESOMAR 1972).
- Lotte Bailyn, "Career and Family Orientations of Husbands and Wives in Relation to Marital Happiness", Human Relations, 23 (1970), pp. 97-114.
- Robert O. Blood and Robert L. Hamblin, "The Effect of the Wife's Employment on Family Power Structure", Social Forces, 36 (May 1958), pp. 347-352.
- Robert O. Blood and Donald Wolfe, Husbands and Wives (New York: The Free Press, 1960).
- Susan Douglas and Patrick Le Maire, "Improving the Quality and Efficiency of Life-Style Research", Proceedings of the XXIII ESOMAR Congress (Hamburg: ESOMAR 1974).
- Abbott L. Ferris, Indicators of Trends in the Social Status of American Women (New York: Russell Sage, 1971).
- Paul E. Green and Donald Tull, Research for Marketing Decisions, third edition, (Englewood Cliffs, N.J.: Prentice Hall, in press).
- Ruth E. Hartley, "Some Implications of Current Changes in Sex and Role Patterns", Merrill-Palmer Quarterly, April 1960, pp. 153-164.
- Janice Hedges and Jeanne Barnett, "Working Women and the Division of Household Tasks", Monthly Labor Review, 95 (April 1972), pp. 9-14.
- David Heer, "Domnance and the Working Wife", Social Forces, 36 (May 1958), pp. 341-347.
- Lois W. Hoffman, (a) "Parental Power Relations and the Division of Household Tasks"; (b) "The Decision to Work", in F. Ivan Nye and Lois Hoffman's The Employed Mother in America, (Chicago: Rand McNally 1963).
- H. Howard and Britt Harris, "A Hierarchical Grouping Routine", I.B.M. 360/65 Fortran Program, University of Pennsylvania Computer Center, 1966.
- Fabian Linden, Women: A Demographic, Social and Economic Presentation, (New York: Conference Board 1973).
- Tony Lunn, Sally Baldwin and Jackie Dickens, "Monitoring Consumer Life-Styles", Admap 8, (November 1973).
- F. Ivan Nye and Lois Hoffman, The Employed Mother in America (Chicago: Rand McNally, 1963).
- Kathryn S. Powell, "Family Variables", Marriage and Family Living, 22 (February 1960), p. 22-35.
- U.S. Dept. of Labor, Report of the President's Commission on Man-Power, 1972.
- "Working Wives", Management Review, (April 1967), pp. 4-21.

SPOUSAL INVOLVEMENT AND EMPATHY IN JOINTLY-RESOLVED AND  
AUTHORITATIVELY-RESOLVED PURCHASE SUBDECISIONS

Alvin C. Burns, University of West Florida

Abstract

Unique characteristics of spousal involvement and empathy in product-feature subdecisions are discussed and depicted on an involvement-empathy graph. A general relationship between spouses' locations on the graph and mutually designated joint subdecisions is described. A similar, but opposite association with recognized authority subdecisions is presented. Implications are discussed.

Problem Setting

The study of husband and wife purchase decision-making has progressed significantly in the past five years. Quite recently, Safilios-Rothschild (1970) criticized work in the area for its naive conceptualizations, methodological inconsistencies, and lack of sensitivity to process. Similar views were expressed by Davis (1970) and Granbois (1971). Largely as a result of these constructive criticisms, there has come about an increase in activity yielding a host of empirical studies, which, although fragmented, has served to restructure thinking about husband and wife decision-making to a considerable degree. For example, the existence of dimensions and stages in spousal purchase decision-making has been documented (Davis, 1970; Davis and Rigaux, 1974); attention has been focused on husband and wife role structure as it varies across product types and decision categories (Davis, 1970; Ferber and Nicosia, 1972; Hempel, 1974; Woodside, 1974); new techniques have been applied to the measurement of perception and influence (Davis, 1971; Ferber and Lee, 1974; Hempel, 1974) and fresh conceptualizations have been suggested (Scott, 1970; Sheth, 1971; Granbois, 1972).

Very recent studies have placed attention on antecedent conditions or determinants of husband and wife role structure. Hempel (1974) and Woodside (1974) have investigated relationships between family role structure and demographics, prior decision-making experiences, and/or psychographics, particularly in the identification of syncretic and spouse-dominant decision areas. While significant associations were found in isolated cases, both authors report considerable unexplained variation and recommend further research. The present study responds by concentrating on predictor variables which are hypothesized to characterize joint and spouse-dominant purchase decision areas. For the purpose of this study, the basic unit of analysis is the point at which spouses settle differences between their first choice preferences for a product feature or subdecision. Justification for concentrating on this phase of the process, i.e., the final choice, is derived from two sources. Logic suggests that it is the juncture wherein a number of forces must interact in order to determine an ultimate choice, if possible. Thus, one might expect to find the maximum impact of determining factors at this point in the husband and wife purchase decision-making process. Moreover, studies (Davis and Rigaux, 1974, for example) have determined that this stage has unique properties.

Use of product-related decision areas is consistent with the trend toward breaking a generic product purchase decision down into more specific subdecisions. Davis

(1970) and Woodside (1974), for example, have discussed the subdecisions involved with automobile as well as rugs and carpet purchases; Hempel (1974) has treated the housing purchase decision in a similar manner, as have Munsinger, Weber, and Hansen (1975). The present study utilizes thirty-nine product feature decisions which represent a sampling of wife-associated, husband-associated, and joint decision areas. Several automobile and television subdecisions were purposely included as part of a larger study. The remaining subdecisions pertain to purchases ranging from housing decisions to clothing decisions. All share a common theme, however, in that the family budget would be affected by the product purchase. (See Table 1 for a listing of the thirty-nine subdecisions.)

Variables Under Study and Rationale

Three husband and wife decision-making variables are discussed in this paper. The three variables are: involvement, empathy, and recognized authority. Each variable has been treated in some respect as a determinant or characteristic of husband and wife preference discrepancy resolution; however, none has been subjected to extensive empirical investigation in the author's knowledge.

Involvement pertains to the importance of the subdecision outcome to the individual spouse. The variable is stated in personal terms (i.e., "How important is it to me ....") and is assumed to be an element contributing to the tenacity with which a spouse defends his or her first choice. Origins of this factor are evident in Morgan (1961) who discusses the strength of a spouse's preferences as derived from the expected subjective utility of the anticipated outcome of the decision. Similar constructs have been suggested as factors affecting husband-wife purchase decision-making outcomes by Brown (1961), Heer (1963) and Granbois (1971). Pollay (1968) incorporates involvement as a process variable in his model of family decision making. In sum, involvement or importance is well established from a theoretical standpoint.

The second variable in this study is a variable which has also been conceptualized as a determinant of a spouse's strength of preference. Empathy pertains to the importance to the individual that his or her spouse's preferences are accounted for in the final product feature choice. While this factor has not received theoretical development commensurate with the involvement variable, empathy has been discussed by Morgan (1961). Clawson (1961) outlines its role in family decision making, while Pollay (1968) has an expanded treatment in his model. He explicitly takes into account the effects of vicariously-derived spousal satisfaction on the decision-making process.

The final variable under study in this paper is recognized authority, defined to be the mutually understood right of one spouse to resolve disagreement between the spouses' first choices. The recognized authority concept differs somewhat from the usual procedures for the identification of spouse-dominant decision areas. Generally, inferences have been drawn from spouses' statements of anticipated influence distributions (for

example: Davis, 1970), recollections (for example: Hempel, 1974) or direct monitoring (Kenkel, 1957). Recognized authority in the current usage pertains to acknowledged decision-making authority in response to the question, "In your own family, who has the right to make the final decision ...?" It is intended to designate mutually acknowledged legitimate authority (or joint) subdecisions rather than inferred ones.

#### Methodology

The data for this study were obtained as part of a larger study on husband and wife preference discrepancy resolution conducted during the spring of 1972. The original study secured the written responses and tape-recorded discussions of 101 married couples who were invited to participate in a behavioral laboratory study concerned with husband and wife purchase decision processes. During the first phase of the study the spouses responded separately to a series of identical questionnaires. At one point they were required to designate which (husband, wife, together, or neither one), in that spouse's opinion, was the most probable mode that would be used to resolve differences between their first choice preferences in each of the thirty-nine product-related subdecision areas. Upon completion of the questionnaires, spouses separately performed a card sorting task which comprised the involvement and empathy instrument. Spouses were required to place each of the thirty-nine product-related subdecision cards in labeled slots corresponding to the spouse's belief with respect to the importance that his or her preferences should be reflected in the couple's designated final choice (involvement). A second set of the randomly ordered decision area cards was also sorted by each participant according to the importance to that individual that his or her spouse's preferences be reflected in the couple's designated final choice (empathy). The sorting responses were assumed to be equivalent to paper-and-pencil responses on an eleven-point scale, and use of the coded cards reduced keypunching time and error.

#### Findings

Standardized scores were computed for each spouse's responses to the thirty-nine product feature subdecisions on the eleven-point importance scale in order to translate them from an absolute to a relative basis. The z-score transformation converts raw scores into standard units of distance away from a spouse's mean score for the thirty-nine decision areas: the transformed values have a mean of zero and unit standard deviation. Direct comparisons between spouses are more meaningful with the use of z-scores as the issue is whether or not the individual is more involved or more empathic with one decision area relative to other decision areas. Thus, personal dispositions are eliminated as comparisons are based on variability.

Spousal group averages and standard deviations were computed for each of the thirty-nine subdecisions. These are presented in Table 1 and Table 2 arranged in descending order of husbands' involvement averages. In the case of the husbands' average involvement levels, the values range from a low of -1.38 to a high of 1.48, while the wives' averages denote a slightly narrower range from -1.27 to 1.23. A different pattern characterizes the empathy averages with the husbands group presenting a range from -1.22 to 1.08 and the wives group averages ranging from -1.36 to .94. It is interesting to note that even though the empathy ranges are identical in width (2.3 standard deviations), the husbands group defines a higher range than the wives group. Multivariate tests of the significance of differences

between and within group means on eleven of the automobile related subdecisions has been reported elsewhere (Burns, 1973; Burns and Granbois, 1975). Significant differences were found at the .001 level between the husbands group means and the wives group means for both involvement and empathy. Within each group and for each variable, significant differences at the .01 level were evident between a number of the subdecisions. It was observed that well-defined multiple levels of involvement and empathy exist for both spousal groups.

Pearson product-moment correlation coefficients were computed for the spousal averages to ascertain the degree of association between the variables. Table 3 reveals that no significant correlation exists between husbands' and wives' average involvement with the thirty-nine subdecisions. Spouses' empathy levels are associated in a moderate negative sense. There is, however, high positive correlation between husbands' involvement with the subdecisions and wives' empathy for husbands' preferences in these subdecisions. Similarly, wives' involvement levels are highly correlated with husbands' empathy levels. Thus it appears that whereas spouses exhibit different responses on the same variable, they are quite compatible across variables, for either spouse's involvement tends to be closely matched with the other spouse's empathy.

Attention was next focused on analysis of the recognized authority responses. Table 4 contains the results of this analysis (the thirty-nine subdecisions are listed in the same order as Tables 1 and 2 to facilitate comparisons). It is important to bring to the reader's attention that the percentages in the table indicate mutually designated response modes, i.e., both spouses indicated the same response. The "other" category refers to a mixed pair such as one spouse indicating "joint" and the other spouse indicating "wife." As expected, the subdecisions present a mix of mutually designated husband-authority, wife-authority, and joint decision areas. Considering only the modal category of these three, there are six decision areas in which the husband is the recognized authority spouse (numbers 4, 5, 9, 15, 19 and 29), eight areas in which the wife is the recognized authority spouse (numbers 30, 32, 34, 35, 36, 37, 38 and 39), and twenty-five areas which are mutually designated as requiring joint resolution (remaining numbers). It is noteworthy that almost all of the husband-authority subdecisions pertain to the family automobile and many of the recognized wife-authority product feature decision areas pertain to homemaking products. The joint decisions, on the other hand, are a more diverse set, as no common theme is discernible at this stage of the analysis. The findings are in large part dictated by the original set of subdecisions, of course, and it is tenuous to generalize at this time. Nevertheless, many of the percentages do support previous findings based upon influence distributions, dominance indices, or observation (see, for example: Kenkel, 1957; Davis, 1971; Woodside, 1972; Davis and Rigaux, 1974; or Hempel, 1974).

An effort was next undertaken to investigate the existence of an association between spousal involvement, empathy, and the designated resolution mode. Visual comparison of Tables 1, 2 and 4 suggests that some relationship does exist between the variables, as the recognized authority mode tends to be matched with low involvement for the nonauthority spouse. At the same time, the authority spouse often indicates low empathy for the nonauthority spouse's preferences. Identification of a clear pattern is hampered, however, by the range of percentages. In an effort to translate involvement and empathy levels into a more graphic form, preliminary analysis entailed plotting husband and wife positions on a two-dimensional graph utilizing involve-

TABLE 1

AVERAGE INVOLVEMENT Z-SCORES\* FOR 39 PRODUCT-FEATURE DECISION AREAS

Subdecision	Husbands		Wives	
	Mean	S.D.	Mean	S.D.
1. Method of financing a new house	1.48	.84	.69	.90
2. Method of financing a new automobile	1.23	.76	.35	.89
3. Price range of a new car	1.01	.57	.33	.85
4. Style of a new suit for husband	.86	.89	-.32	.76
5. Type of brakes in a new car	.68	.92	-.76	.86
6. Size of a new automobile	.66	.67	-.16	.81
7. When next auto will be purchased	.60	.65	-.16	.81
8. Make of a new car	.60	.75	-.49	.74
9. Type of transmission on a new car	.51	.80	-.63	.91
10. Number of rooms in a new home	.48	.76	1.23	.69
11. How much to spend on summer vacation	.46	.79	.38	.82
12. Model of a new automobile	.43	.71	-.44	.71
13. Price range of a new TV	.37	.67	.11	.75
14. Body style of a new automobile	.31	.68	-.58	.73
15. Place of purchase for a new automobile	.25	.82	-.92	.74
16. Where summer vacation will be spent	.22	.81	.65	.89
17. How much to spend on stereo record player	.21	.85	.06	.67
18. Amount to spend on new living room furniture	.20	.85	.60	.64
19. Engine horsepower of a new car	.17	.80	-1.27	.58
20. Color or black and white picture tube	.07	.92	-.05	.84
21. Brand of new TV	.02	.72	-.56	.75
22. When to purchase a television set	-.20	.65	-.35	.62
23. Picture tube size	-.26	.80	-.69	.65
24. Exterior color of new auto	-.27	.67	-.54	.73
25. When living room furniture should be replaced	-.33	.73	.43	.67
26. Where a new TV will be purchased	-.34	.77	-.86	.74
27. When a refrigerator will be purchased	-.36	.67	.29	.72
28. Seat upholstery for new car	-.42	.82	-.60	.67
29. Type of radio in new auto	-.43	.80	-1.06	.65
30. Brand of new washer dryer	-.55	.72	.40	.78
31. Cabinet style of new TV	-.57	.66	.07	.82
32. Freezer size of new refrigerator	-.64	.69	.73	.90
33. Where to buy living room furniture	-.68	.70	-.03	.88
34. Style of a new dining room set	-.71	.70	.80	.61
35. Color and fabric of living room set	-.75	.66	1.06	.59
36. Style of new cocktail dress	-.93	.74	.73	.80
37. Length of new cocktail dress	-.95	.71	.57	.98
38. Capacity of new washer-dryer	-1.07	.63	.78	.70
39. Color of a new oven	-1.38	.42	.21	.86

\*Ordered by husbands' involvement

ment and empathy as independent axes. Such a presentation is defensible inasmuch as the association between the two variables for the wives' averages was not significantly different from zero correlation at the .01 level. The husbands' average levels, one will recall, did yield a correlation coefficient which was above the cutoff value at the .01 level. The association was found to be moderate; thus, distortion for husbands' positions is slight. It is reported elsewhere that eleven decision areas support the hypothesis of independence of the two variates for both spousal groups (Burns, 1973; Burns and Granbois, 1975).

The spousal positions are presented in Figure 1 which portrays several of the findings reported earlier. Each spousal group average coordinate is located on the dimensions and connected to the corresponding (other) spousal group coordinate with a line. The numbers beside pertain to the numbers assigned to each subdecision in Tables 1, 2 and 4. Almost without exception the lines slope down and to the right which is the arrangement suggested by

the high positive correlations and range comparisons described in the previous discussion. Plainly stated, one spouse tends to exhibit greater involvement than the other spouse in any given subdecision, but the less involved spouse is more empathic. This pattern holds regardless of whether the husband or the wife is more involved.

Presumably the less involved, more empathic spouse assumes a passive role in the decision-making process. However, while Figure 1 depicts involvement and empathy configurations, it is difficult to make direct comparisons between these dispositions and the recognized preference discrepancy resolution mode for a particular subdecision. Interpretation is further confused by some lines which cut across quadrants and others which are contained within a single quadrant. It is apparent that some subdecisions are quite similar with respect to both their perceived importance and the degree to which the spouse's preferences are important. For example, the



TABLE 2

## AVERAGE EMPATHY Z-SCORES\* FOR 39 PRODUCT-FEATURE DECISION AREAS

Subdecision	Husbands		Wives	
	Mean	S.D.	Mean	S.D.
1. Method of financing a new house	.16	.99	.94	.63
2. Method of financing a new automobile	-.12	.98	.80	.65
3. Price range of a new car	-.02	.77	.75	.60
4. Style of a new suit for husband	-.57	1.00	.70	.85
5. Type of brakes in a new car	-1.20	1.14	.65	.66
6. Size of a new automobile	-.08	.77	.40	.62
7. When next auto will be purchased	-.25	.80	.54	.65
8. Make of a new car	-.28	.56	.57	.68
9. Type of transmission on a new car	-.62	1.01	.62	.73
10. Number of rooms in a new home	.97	.62	.54	.83
11. How much to spend on summer vacation	.23	.66	.58	.70
12. Model of a new automobile	-.36	.70	.50	.66
13. Price range of a new TV	-.20	.77	.39	.63
14. Body style of a new automobile	-.30	.70	.38	.82
15. Place of purchase for a new automobile	-.83	.86	.24	.76
16. Where summer vacation will be spent	.67	.76	.59	.73
17. How much to spend on stereo record player	-.31	.75	.17	.67
18. Amount to spend on new living room furniture	.27	.78	.21	.70
19. Engine horsepower of a new car	-1.22	.88	.66	.71
20. Color or black and white picture tube	-.10	.77	-.11	.77
21. Brand of new TV	-.61	.74	-.05	.83
22. When to purchase a television set	-.28	.77	.05	.65
23. Picture tube size	-.38	.70	-.02	.69
24. Exterior color of new auto	-.08	.63	-.19	.86
25. When living room furniture should be replaced	.40	.78	-.24	.74
26. Where a new TV will be purchased	-.79	.70	-.34	.84
27. When a refrigerator will be purchased	.27	.80	-.23	.78
28. Seat upholstery for new car	-.19	.63	-.29	.79
29. Type of radio in new auto	-.77	.78	-.05	.74
30. Brand of new washer dryer	.43	.75	-.82	.90
31. Cabinet style of new TV	.21	.71	-.50	.70
32. Freezer size of new refrigerator	.69	.87	-.88	.86
33. Where to buy living room furniture	.01	.93	-.65	.79
34. Style of a new dining room set	.81	.76	-.48	.90
35. Color and fabric of living room set	.85	.77	-.69	.78
36. Style of new cocktail dress	1.08	.88	-.95	1.06
37. Length of new cocktail dress	.90	.76	-1.14	.96
38. Capacity of new washer-dryer	.87	.97	-1.28	.88
39. Color of a new oven	.74	.67	-1.36	.73

\* Ordered by husbands' involvement

TABLE 3

CORRELATION MATRIX FOR HUSBANDS' AND WIVES'  
AVERAGE INVOLVEMENT AND EMPATHY LEVELS

	Husband Involvement	Husband Empathy	Wife Involvement	Wife Empathy
Husband Involvement	1.00	-.47*	-.18	.95*
Husband Empathy		1.00	.91*	-.56*
Wife Involvement			1.00	-.31
Wife Empathy				1.00

\*Correlation significant at .01 level

TABLE 4

## MUTUALLY DESIGNATED\* RESOLUTION MODES FOR THE SUBDECISION

Subdecision	Resolution Mode (Percentages)			
	Husband	Wife	Joint	Other
1. Method of financing a new house	14	0	56	30
2. Method of financing a new auto	20	0	46	35
3. Price range of a new car	9	0	53	39
4. Style of new suit for husband	63	1	12	24
5. Type of brakes in a new car	75	0	0	25
6. Size of a new auto	6	0	61	33
7. When next auto will be purchased	15	0	42	44
8. Make of a new car	27	0	29	45
9. Type of transmission on new car	47	0	14	40
10. Number of rooms in a new home	0	5	74	21
11. How much to spend on summer vacation	3	0	74	23
12. Model of a new auto	22	0	34	45
13. Price range for a new TV	1	0	63	36
14. Body style of a new auto	11	0	47	43
15. Place of purchase for a new auto	32	0	12	56
16. Where summer vacation will be spent	1	0	91	8
17. How much to spend on stereo record player	6	0	65	29
18. Amount to spend on new living room furniture	0	2	60	38
19. Engine horsepower of a new car	87	0	1	12
20. Color or black and white picture tube	3	1	70	26
21. Brand of a new television	12	1	50	38
22. When to purchase a new television set	4	0	74	22
23. Picture tube size of a new TV	2	0	63	35
24. Basic exterior color of a new auto	1	1	61	37
25. When living room furniture should be replaced	1	13	37	50
26. Where a new Tv will be purchased	7	0	53	41
27. When a refrigerator will be purchased	0	5	50	45
28. Seat upholstery for new car	3	4	32	61
29. Type of radio, if any, in new auto	32	0	12	56
30. Brand of a new washer dryer	0	43	14	44
31. Cabinet style of new TV	0	7	48	46
32. Freezer size of new refrigerator	0	48	14	39
33. Where to buy living room furniture	0	14	34	53
34. Style of new dining room set	0	42	16	43
35. Color and fabric of new living room furniture	0	56	10	34
36. Style of new cocktail dress	0	75	3	22
37. Length of new cocktail dress	0	76	4	20
38. Capacity of a new washer-dryer	0	76	2	23
39. Color of a new oven	0	82	1	17

\*Both spouses indicated the same mode; e.g., both indicated "Husband".

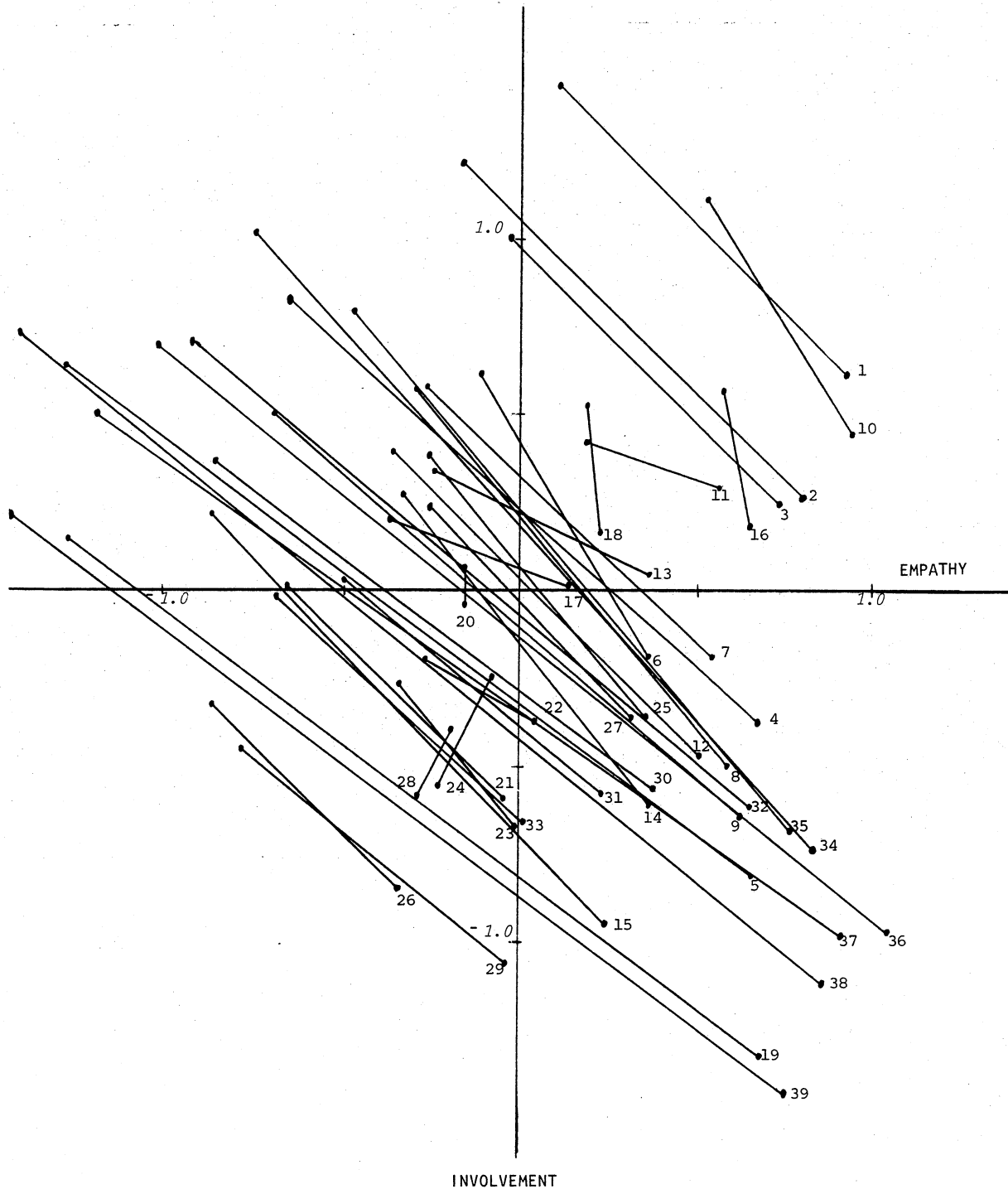
"How much to spend on your next summer vacation" (number 11) decision is located approximately in the same place by both spouses in the positive involvement-positive empathy quadrant. Other subdecisions represent quite dissimilar positions: the "Capacity of a new washer-dryer" (number 38) is a good example. The decision was made to use the Euclidean distance measure (squared) as an index of the proximity of spousal positions on the two dimensions. This distance measure is in common use and is applicable even if the variables are correlated (Green and Tull, 1970).

Squared Euclidean distances were computed between the spousal group average z-scores for each of the thirty-nine subdecision areas. Correlation coefficients were computed to assess the association between the squared distances and the percent of the sample which mutually designated a joint decision and the percent which indicated an authority spouse (either one). These coefficients were found to be  $-.83$  and  $.95$  respectively, re-

vealing two strong and opposite linear relationships significant at the .01 level. Both of these associations are readily apparent in Figures 2a and 2b. In the instance of the comparison of the distance measure with the percentage of the sample which mutually designated the preference discrepancy resolution mode as joint, the likelihood of joint resolution diminishes rapidly as the distance between the spousal positions increases. Alternatively stated, one can see that the more similar the spousal positions on the two variables, the more likely the resolution of any preference discrepancies will take the form of joint participation. Thus, regardless of whether both spouses are greatly involved (empathic) or minimally involved (empathic) with the outcome of a particular product-related decision, joint decision making will tend to be reported as long as they occupy similar empathy (involvement) dispositions. Conversely, mutually designated joint preference discrepancy resolution is indicative of similar dispositions, although the precise location on the dimensions

FIGURE 1

HUSBAND AND WIFE POSITIONS ON INVOLVEMENT AND EMPATHY DIMENSIONS



NOTE: Numbers correspond to those in Table 1.

is not suggested. The diagram also points out that the relationship is probably nonlinear rather than linear.

The recognized authority percentages and associated distances yield a pattern which complements the one found in the previous comparison. While the points are somewhat more scattered, it is nonetheless clear that the likelihood of a mutually agreed recognized authority

figure increases directly with the distance between the spousal involvement and empathy positions. Those product feature decisions which are mutually designated as under the legitimate authority of one spouse (either one) are characterized by large differences between each spouse's perception of the importance of the outcome and the degree of vicarious satisfaction anticipated. Considering other findings in this study, it is

FIGURE 2a

COMPARISON OF DISTANCES AND DESIGNATED JOINT SUBDECISIONS

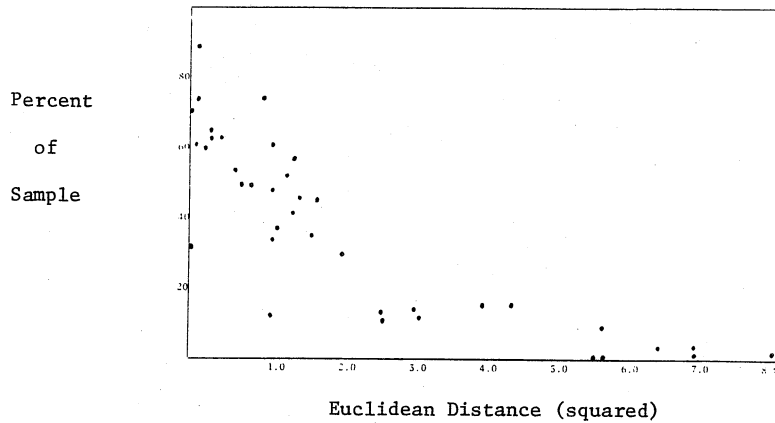
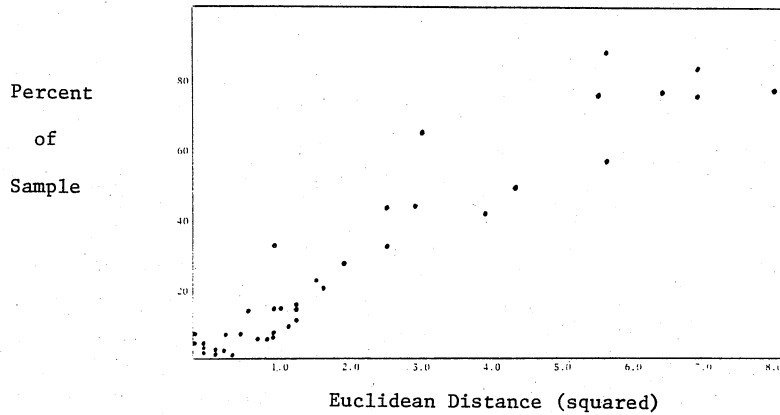


FIGURE 2b

COMPARISON OF DISTANCES AND DESIGNATED AUTHORITY SUBDECISIONS



also proper to expect that the authority spouse will be more highly involved and less empathic than the non-authority spouse.

Discussion

It is necessary to preface this section with a description of the sample as it was secured with a nonrandom procedure. Consequently, knowledge of its demographic characteristics will serve to temper comments. The sample conforms to what one might expect would result from a voluntary participation in a study conducted in a mid-western city with a large university population. The average husband age was thirty-nine years, while the average wife age was thirty-seven years. These means exhibited large standard deviations of 10 and 9.5 years respectively. Both spouses were typically college graduates, and the average length of marriage was fifteen and one-half years. Married couples with husbands who were full-time students were not allowed to participate in the study. Substantial variance was found in the sample demographics; nevertheless, it is evident that both spouses were well-educated, mature, financially secure, and had experienced several years of family purchase decision making together.

Keeping these factors in mind, a striking aspect of the

findings which deserves discussion is the high degree of compatibility exhibited by spousal involvement with and empathy for the several product-feature areas. Speaking strictly from a speculative standpoint, insight with respect to the dynamics and results of husband and wife preference discrepancy resolution is perhaps evident. Specifically, when one spouse indicates high involvement with a subdecision, the other appears to realize this state and reciprocates with a greater willingness to allow that spouse's preferences to shape the final outcome. Thus, serious conflicts are avoided and the outcome is satisfactory to each spouse but for different reasons. Assuming the outcome of preference discrepancy resolution is dictated by the two variables, the highly involved spouse derives satisfaction by virtue of the fact that his or her preferences are reflected in the outcome, while the highly empathic spouse derives gratification vicariously. The findings suggest that those decision areas which are within the legitimate decision-making authority of a particular spouse are readily acknowledged and the nonauthority spouse concedes the decision graciously to that spouse. Research into these questions seems justified given the strong associations found in this study. Also, the findings offer support at the descriptive level for Pollay's (1968) husband and wife decision-making conceptualization which utilizes involvement and empathy as dynamic, process variables.

Implications as to the nature of jointly resolved product feature decision areas are also provided by the findings. It is most significant that the proximity of spousal positions on the two dimensions is directly associated with the probability of a subdecision being mutually acknowledged as syncratic. Intuitively, one might suspect that the high-involvement, high-empathy quadrant would be the locus of joint decisions, but this hypothesis is refuted by the fact that joint decisions are found in both the low-low quadrant and across quadrants. An intriguing question which should be addressed in future research is the effect of a subdecision's joint quadrant location on the exact joint resolution mode. For example, are certain dispositions indicative of compromise, concession, impasse, or some other mode? Furthermore, relative differences between husband and wife locations are another factor which may explain or describe decision-making characteristics. Inasmuch as both variables allude to the tenacity with which a spouse maintains his or her preferences during decision-making, the degree of influence shared in joint decision making is logically associated with them. Consequently, it is appropriate to investigate influence distributions in light of relative involvement and/or empathy differences. Attention should be given to these questions in future research with the ultimate goal of determining the predictive value of spousal involvement and empathy within a particular subdecision. The nature of conflict resolution strategies and outcomes may become more clear if relationships between these variables are documented.

A separate but equally important implication of the study's findings concerns the prospect of vicarious satisfaction within the husband and wife decision making process. While the use of recall and perceptions has been thoroughly discussed in the literature (Ferber, 1955, Safiliou-Rothschild, 1970; Davis, 1971), it is important to comment on the interaction which might exist between a spouse's statement of influence in a particular decision or decision sequence and the vicarious satisfaction which that spouse may have gained from the outcome(s) of the decision(s). Consider, for example, an instance in which the authority spouse selects the nonauthority spouse's first preference. If this state arose, the latter might be inclined to report greater participation in the decision regardless of the actual degree of participation. The effects of independently-resolved decisions are similarly implied by the findings, for it would seem that the uninvolved spouse may still derive gratification from the process even when the other acts autonomously or without prior consultation. At this point it seems appropriate to hypothesize that there may be a difference between process-derived satisfaction from outcome-originated satisfaction.

While additional questions are evident in the findings, it is premature to discuss them at length. Instead, the conclusive aspects of the study will serve to terminate discussion. The findings are reasonably conclusive in the sense that they lend support for the division of generic product into product-related subdecisions as advocated and practiced by Davis (1971), Hempel (1974) and Woodside (1974). It is significant that the methodology and variables in this study are different from those used by these authors yet the conclusion is identical. The findings are also reasonably conclusive in that involvement and empathy appear to be productive dimensions with which to describe spousal dispositions. In terms of spouses' predictions, joint decisions are identifiable. In the instance of joint decisions, spouses tend to exhibit similar positions; recognized authority decisions, on the other hand, are associated with dissimilar positions. Future research in this area seems worthy and advisable.

## References

- George H. Brown, "The Automobile Buying Decision Within the Family," in Nelson N. Foote, Household Decision-Making (New York: New York University Press, 1961).
- Alvin C. Burns, "A Study of Husband and Wife Preference Discrepancy Resolution," unpublished doctoral dissertation, Graduate School of Business, Indiana University, (1972).
- Alvin C. Burns and Donald H. Granbois, "Process Variables in Husband and Wife Preference Discrepancy Resolution," unpublished paper, (1975).
- Joseph C. Clawson, "Family Composition, Motivation, and Buying Decisions," in Nelson N. Foote, Household Decision-Making (New York: New York University Press, 1961).
- H. L. Davis, "Dimensions of Marital Roles in Consumer Decision-Making," Journal of Marketing Research, 7 (May, 1970), 168-77.
- \_\_\_\_\_, "Measurement of Husband-Wife Influence in Consumer Purchase Decisions," Journal of Marketing Research, 7(August, 1971), 305-12.
- H. L. Davis and B. P. Rigaux, "Perception of Marital Roles in Decision Processes," Journal of Consumer Research, 1(1974), 51-62.
- Robert Ferber, "On the Reliability of Purchase Influence Studies," Journal of Marketing, 19(January, 1955), 225-232.
- Robert Ferber and L. C. Lee, "Husband-Wife Influence in Family Purchasing Behavior," Journal of Consumer Research, 1(1974), 43-50.
- Robert Ferber and F. M. Nicosia, "Newly Married Couples and Their Asset Accumulation Decisions," in Human Behavior in Economic Affairs (Amsterdam: North Holland Publishing Company, 1972).
- Donald H. Granbois, "A Multilevel Approach to Family Role Structure Research," Proceedings. Second Annual Conference Association for Consumer Research, (1971), 91-107.
- \_\_\_\_\_, "Decision Processes for Major Durable Goods," in George Fisk, New Essays in Marketing Theory (Boston: Allyn and Bacon, 1972).
- Paul E. Green and Donald S. Tull, Research for Marketing Decisions (Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1970).
- Donald J. Hempel, "Family Buying Decisions: A Cross Cultural Perspective," Journal of Marketing Research, 11(1974), 295-302.
- David M. Heer, "The Measurement and Bases of Family Power: An Overview," Marriage and Family Living, (1963), 133-139.
- William F. Kenkel, "Influence Differentiation in Family Decision-Making," Sociology and Social Research, 42 (September-October, 1957), 18-25.
- James N. Morgan, "Household Decision-Making," in Nelson Foote, Household Decision-Making (New York: New York University Press, 1961).

Gary M. Munsinger, Jean E. Weber, and Richard W. Hansen, "Joint Home Purchasing Decisions by Husbands and Wives," Journal of Consumer Research, 1(1975), 60-66.

Richard W. Pollay, "A Model of Family Decision-Making," British Journal of Marketing, (Autumn, 1968), 206-216.

Constantina Safilios-Rothschild, "The Study of Family Power Structure: A Review 1960-1969," Journal of Marriage and the Family, (November, 1970), 539-552.

R. A. Scott, "Husband-Wife Interaction in a Household Purchase Decision," Southern Journal of Business, 5 (July, 1970), 218-25.

J. N. Sheth, "A Theory of Family Buying Decisions," in J. N. Sheth, Models of Buyer Behavior (New York: Harper and Row, 1974).

A. G. Woodside, "Dominance and Conflict in Family Purchasing Decisions," Proceedings, Third Annual Conference Association of Consumer Research, (1972).

\_\_\_\_\_, "Effects of Prior Decision-Making, Demographics, and Psychographics on Marital Roles for Purchasing Durables," Proceedings, Fifth Annual Conference Association for Consumer Research, (1974).

## HOW WILL CONSUMER EDUCATION AFFECT CONSUMER BEHAVIOR?

Paul N. Bloom, University of Maryland\*

### Abstract

The recent upsurge of interest in consumer education can be expected to lead, in the long-run, to significant changes in consumer behavior. This paper contains a brief description of existing consumer education programs and a discussion of several hypotheses about how programs of this type could affect consumer behavior.

Interest in consumer education is growing rapidly. Within the last few months, a conference on consumer education has been held at the White House, an Office of Consumers' Education has been established within the Department of Health, Education and Welfare, and courses in consumer education have been made mandatory for all high school students in several states. This growth of interest in consumer education can be attributed to the recent depressed economic situation and, to some extent, the disappointing results achieved by certain consumer information programs (see Day and Brandt, 1974; Jacoby, Speller, and Kohn, 1974; Isakson and Maurizi, 1973). Many individuals have come to believe that consumer education programs are needed to help consumers deal with inflation and energy shortages and to teach consumers how to use and benefit from "Truth in Packaging," "Truth in Lending," and "Unit Pricing" disclosures.

If interest in consumer education continues to grow, and if more and better consumer education programs follow, then significant longrun changes can be expected to occur in the behavior of consumers. The consumers of tomorrow will have gone through a very different consumer socialization process than the consumers of today, and this could lead to vastly different consumer expectations, attitudes, preferences, and shopping habits. The purpose of this paper, therefore, is to provide an overview of the consumer education programs currently in operation in the United States and to present several hypotheses about how programs of this type could affect consumer behavior in the future. Before this material can be presented, however, it is necessary to clarify what interpretation is given in this paper to the terms "consumer education" and "consumer education program."

### A Definition of Consumer Education

In this paper, consumer education is viewed as the process by which people learn the workings of the marketplace so that they can improve their ability to act as purchasers or consumers of those products and

services they deem most likely to enhance their well being (see Willett, 1974; Seitz, 1972, p. 199). Consumer education is therefore treated as being rather different than consumer information -- something with which it is often confused. Consumer education is considered to be a learning process which people go through which, of course, cannot be readily observed or heard. Consumer information, on the other hand, is clearly something which can be observed or heard.

A consumer education program is viewed in this paper as any organized activity which has as one of its major ultimate goals the advancement of the process of consumer education among some segment of consumers. This definition permits programs which certain people in the consumer education field might label as consumer information programs -- because no teaching in a formal educational environment is taking place -- to be viewed as consumer education programs. For example, a program which develops consumer information pieces such as buying guides or curriculum guides for use in high school consumer education courses is considered a consumer education program. Similarly, a program in which advertising messages are designed and distributed that instruct consumers in how to shop for certain categories of products is considered a consumer education program. Both of these programs are designed to help advance the process of consumer education. However, programs which provide only factual information about specific product offerings, from which consumers can learn little that can be used in a variety of buying situations (e.g., unit pricing or "Truth in Lending" programs), are considered merely consumer information programs.

### Current Consumer Education Programs

There are now a very large number of consumer education programs in operation in the United States. A major study conducted at Purdue University in 1969 (Uhl, 1970) identified and examined more than 500 of these programs, and many more have been organized since that time. These programs are administered by a wide variety of public and private organizations, and are designed to help consumers of all age brackets, social classes, and ethnic backgrounds. A brief review follows of the consumer education programs administered by Federal, state, and local governments and by private non-profit and profit-making organizations.

The Federal government has a host of agencies administering consumer education programs. Some of these agencies set up and run model programs for other organizations to copy. Other agencies produce and distribute consumer information materials which are used by the media to help design consumer education messages or by teachers to help instruct consumer education classes. Still other agencies are involved in stimulating additional interest in consumer education among educators, businessmen, and the general public through conducting workshops, symposiums, and seminars. Finally, there are a few agencies which distribute Federal funds to consumer education programs.

The Federal agencies which are most actively involved with consumer education include the Office of Consumer Affairs (within the Department of Health, Education,

\*This research was funded in part by a grant from the General Research Board of the University of Maryland. The author would like to thank Philip G. Kuehl for his helpful suggestions.

and Welfare), the Department of Agriculture, the Office of Economic Opportunity, the Consumer Product Safety Commission, the Federal Trade Commission, and the newly formed Office of Consumers' Education (within the Office of Education of HEW). The Office of Consumer Affairs, for example, publishes and distributes a large variety of materials related to consumer education including curriculum guides, bibliographies, directories, and a newsletter called Consumer News. This agency also sets up and runs model consumer education programs, including one for the residents of the Indian Pueblos of New Mexico, and it is continually arranging seminars and meetings to stimulate interest in consumer education.

The new Office of Consumers' Education also deserves special mention. This agency was established in 1975 to act as a source of funds for consumer education research and programs. Approximately \$3 million per year has been budgeted for this agency during fiscal years 1976 and 1977. This is the first Federal agency that has been given the right to award funds specifically for consumer education research. In the past, funds for this purpose have had to come from monies appropriated for vocational education and other areas. It should be noted that President Ford supported the appropriation of \$3 million per year to this agency, and that he has also demonstrated his support for consumer education by holding a one-day conference on the subject at the White House on March 11, 1975.

State governments tend to get most involved with consumer education through their control over public school systems. However, many state governments also run consumer education programs to help the poor, disadvantaged, or other segments of their populations such as ex-convicts (Goetting, 1974). In the schools, at least six states (Illinois, Oregon, Wisconsin, Florida, Kentucky, and Hawaii) have decided to make consumer education a required course (Brooks, 1973), and many other states have recommended that consumer education be included in the high school curriculum. Thus far, consumer education courses have tended to focus only on "good buymanship" (Scherf, 1974; Uhl, 1970) -- how to maximize current family consumption by stretching dollars wisely and how to maximize future family consumption through wise savings and investments -- but topics such as consumer protection, consumer redress, and "values clarification" have received increasing attention.

Typical of the more innovative programs being conducted in the public schools is one being run in Irving, Texas. Twelfth-grade students are asked to simulate the actions they would have to take if they were to become totally independent from their families upon high-school graduation. They are required to search the want ads for a suitable job (for which they are qualified) and then plan a life style for themselves supported by the take-home pay of that job. They must find an apartment (with furniture), purchase an appropriate wardrobe, budget for cleaning and laundry expenses, open charge accounts, plan and conduct food-shopping trips, and pay for and maintain automobiles. The course has helped many students achieve a well-planned independent life, while it has helped others see the need for further schooling or for improving their relationships with their families (Mahan, 1972).

The consumer education programs administered by local governments are numerous and diverse. There are consumer affairs offices and other agencies in many localities which devote a great deal of time, effort and resources to consumer education programs. Some programs are designed to help consumers with special needs --

such as the poor, Spanish-speaking, or elderly -- and others are more general in nature and, in effect, attempt to help a broad spectrum of individuals. Information about local government programs can be found in two studies by the Office of Consumer Affairs (1973;1974).

Several private, non-profit organizations such as Consumers Union, the Council of Better Business Bureaus, and the American Council on Consumer Interests are extremely active in consumer education. These organizations conduct consumer education classes, publish consumer magazines and newsletters, produce consumer messages for the media, and provide consumer education teachers with valuable teaching materials. They also provide guest speakers for classes and meetings, conduct conferences where consumer education teachers can exchange ideas, and sponsor workshops.

The Consumer Services Program of the Baltimore Urban League is an example of a highly successful consumer education program run by a non-profit organization. This program involves the teaching of six-week mini-courses in consumer education to community groups and clubs (e.g., Golden Agers, school parent groups, church groups, etc.). Participants are awarded certificates upon completion of these courses. Besides the courses, the program also has its staff members working on a one-to-one basis with approximately three to five families, helping them set up their monthly budgets and plan purchases of food, clothing, and other household items. Those involved with the Baltimore Urban League program attribute much of its success to the strategy of working with incentives (i.e., the certificates) and established groups (rather than newly-formed ones) (Johnson, 1975).

Consumer education programs are also now being run by many business firms. Profit-making organizations are providing consumer education teachers with a wealth of booklets, transparencies, film strips, product samples, and guest speakers. They are sponsoring consumer education media messages (e.g., Giant Foods (see Peterson, 1974)) and are even conducting in-service courses for teachers of consumer education (e.g., Montgomery Ward). Some firms, like the publishers of Changing Times magazine, have given considerable support to consumer education programs because this support has produced immediate returns for them (i.e., more subscriptions). Other firms, like the major department store chains (e.g., Sears, J. C. Penney), have made substantial investments in consumer education in anticipation of long-run returns. They apparently feel that support of consumer education can lead in the long-run to more consumer satisfaction with their products and, perhaps, less government interference with their operations.

#### Hypothesized Changes in Consumer Behavior

Although many of the current consumer education programs have been in existence for quite some time, there has not been any research reported on how these programs have affected consumer behavior. Clearly, cross-sectional and longitudinal studies should be conducted to try to measure how much various consumer education programs have changed consumer attitudes, preferences, habits, and so on. It should be recognized, however, that research in this area will be filled with many obstacles. It will be difficult to determine whether a change in consumer behavior has been caused by a consumer education program or by other confounding or intervening variables. Moreover, with some programs, it may take long periods of time before program-induced changes in consumer behavior can be detected -- the learning process can be very slow.



A group of hypotheses that could be tested when re-searching the effects of consumer education programs are presented and discussed below. These hypotheses describe the long-run changes that the author expects to occur in the behavior of most consumers as a result of continued exposure to a variety of consumer education programs. In the absence of any past research from which inferences could be drawn, it was necessary to rely upon deductive reasoning to develop these hypotheses.

Hypothesis 1: Consumers will express their wants and needs to sellers more explicitly and more frequently.

Exposure to consumer education programs should allow consumers to obtain a better understanding of what the marketplace has to offer them in terms of product variations and services. In addition, exposure to consumer education courses which cover areas like "values clarification" or "life adjustment" -- such as the course offered at a Portland, Oregon high school in which students simulate real-life adult experiences including getting married, buying a home and car, and getting a divorce ("Divorce Course", 1974) -- should help many consumers obtain a better understanding of what they want and need from life and from the marketplace. An improved understanding of both what is available and what is wanted from the marketplace should result in an improved ability among consumers to state explicitly what they want and need to sellers in the form of unsolicited letters, responses to marketing research questions, or other methods of communication. Suggestion letters from consumers should become more helpful to sellers and questionnaire responses of consumers should become less ambiguous. Moreover, consumers should become more likely to communicate their wants and needs to sellers as they learn, through consumer education, more about how consumer feedback can influence the decisions of these sellers. The number of suggestion letters and questionnaire responses received by sellers should increase.

Hypothesis 2: Consumers will seek out and use larger amounts of information to help them make price, quality, and service comparisons before making a purchase.

By teaching consumers where consumer information can be found, how it can be used, and what the benefits are of using it, consumer education programs should tend to lower the perceived costs (in lost time and energy) and increase the perceived benefits of searching for and using consumer information. This, in turn, should lead consumers to seek out and use larger quantities of information from consumer magazines or advertising, unit pricing, open dating, and nutritional content disclosures before making a purchase. Consumers must be expected to "buy" more of a commodity (information) which they perceive to be lower in price and higher in quality than it was previously.

Hypothesis 3: Consumers will make a progressively larger proportion of their purchases from sellers who offer relatively large amounts of easily-acquired information about their products and services.

A seller who provides consumers with a considerable amount of information about his product(s) or service(s) (e.g., ingredients, uses, warranties, unit price, advantages, etc.), through either advertising, label disclosures, or salespersons, will save consumers time and energy that could be lost in searching for this information. If consumers (as a consequence of consumer education) begin to seek out and use

larger amounts of information before making purchases -- as hypothesized above -- then they should become progressively more appreciative of sellers who can save them information search costs. As consumers develop more favorable attitudes toward information-providing sellers, the likelihood that they will actually prefer these sellers over others will increase. Thus, consumers can be expected to make an increasing proportion of their purchases from information-providing sellers. They will not buy inferior or overpriced products from these sellers, but will buy the products of these sellers with increasing frequency in those cases where all products competing in a market are perceived to have otherwise similar attributes.

Hypothesis 4: Consumers will make a progressively larger proportion of their purchases from sellers who conduct relatively large consumer education programs.

Over time, large numbers of people can be expected to be exposed to consumer education programs. As more people benefit from their experience with these programs, more are likely to become supporters of consumer education. One way these people might show their support for consumer education is by becoming loyal patrons of sellers whose consumer education programs they personally found to be beneficial. Thus, consumers can be expected to make an increasing proportion of their purchases from sellers who conduct relatively large consumer education programs. Evidence that this trend may have already begun can possibly be found in the rapidly improving profit figures of Giant Food, Inc., a firm which is very active in consumer education (Levy, 1975).

Hypothesis 5: Consumers will buy less products that are potentially harmful to their own health or to the health of others.

Consumer education programs often teach consumers about nutrition, product safety, health care, and environmental issues. As more people are exposed to these programs, one can expect a decline in purchase of products which could produce health problems. Consumers will become less likely to buy products such as cigarettes, alcoholic beverages, sugar-laden foods, low-mileage automobiles, or aerosol sprays, and they will become more likely to buy products such as low-cholesterol foods, automobiles with safety and pollution-control features, or dental floss. It should be noted, however, that it will be necessary to reinforce constantly an inclination that might develop among consumers to buy safer, healthier products. If messages about nutrition, product safety, health care, and pollution are not constantly received by consumers, they may forget what they have learned in these areas. For example, many consumers seem to have forgotten what they learned about the dangers of cigarette smoking in the absence of frequent anti-smoking messages. The volume of these messages was reduced significantly following the banning from the broadcast media of cigarette advertising in 1971. Since that time, cigarette sales have increased steadily, reversing several years of decline.

Hypothesis 6: Consumers will more actively seek remedies, redress, or restitution when dissatisfied with a product, service, or marketing practice.

Through exposure to consumer education programs, consumers should learn where to go when they are dissatisfied with an experience they have had in the marketplace. They should learn which agencies,

organizations, and individuals can help them if they have been deceived or have found a practice they would like to see discontinued. They should also learn how consumers can achieve highly favorable results by filing complaint letters and lawsuits. This knowledge should lead consumers to become much more active in filing complaints and lawsuits against business firms.

Hypothesis 7: Consumers will become more active participants in the debates over legislation that could affect the workings of the marketplace.

Consumer education programs should teach consumers about the important role that government plays in our economy. They should learn about the laws which exist to protect and help consumers and about the legislative history of these laws. They should discover how important it is, from their point of view, to have their voices heard, along with those of representatives of businesses and government, during debates over legislation that could affect their ability to obtain what they want and need from the marketplace. Consumers should therefore become more active in writing to legislators, lobbying for or against legislation, testifying before legislative committees, and supporting politically-active consumer organizations.

#### A Few Comments

The possibility that the changes hypothesized above might occur in the very near future should be recognized. Business firms could respond to a changing consumer population, and the changing practices of competing firms, by trying to outdo one another in creating progressive, extensive consumer education programs. At the same time, legislators might respond to the increased political activity of consumers by starting and funding new consumer education programs and by requiring consumer education courses for all public school students. The amount of exposure consumers would receive to consumer education programs could therefore increase rapidly, leading to the hypothesized changes in consumer behavior at an earlier point in time.

No matter when they occur, the changes hypothesized above will probably occur more rapidly among middle and upper class consumers, younger consumers, and consumers living in or near large metropolitan areas. These groups are more likely to be exposed to consumer education courses, messages, and materials through school programs, the media, and friends. They are also more likely to learn about the workings of the marketplace at a fast pace. Changes in the behavior of poor, elderly, and rural consumers will probably come much more slowly. Consumer educators will have to do a considerable amount of research to find the best strategies for "marketing" consumer education to these segments.

It should be noted that nothing has been mentioned thus far about the consumer of the future paying less attention to advertising or buying more lower-priced, private-label brands -- two results that many people might expect from continued exposure to consumer education programs. These changes have not been hypothesized to occur because it seems equally plausible to expect opposite results. Consumers could pay greater attention to advertising as part of their efforts to acquire more information before making purchases. Similarly, consumers could buy less private label brands because of an increased desire to buy familiar, widely-available products

about which they have a great deal of favorable information.

Finally, it should be mentioned that if consumer behavior changes in the directions hypothesized above as a consequence of consumer education, it will not necessarily be totally beneficial for consumers. If consumers buy less cigarettes, alcoholic beverages, and low-mileage automobiles, then unemployment could increase. If consumers tie up the courts with lawsuits against sellers, then crime in the streets could increase. And if consumers lobby for more consumer protection legislation, then prices charged by sellers could increase (to cover the cost of compliance with new laws). In short, consumer education might unintentionally hurt the interests of many consumers.

#### Conclusions

In the future, consumers will go through a different consumer socialization process than consumers do today. They will spend more time in consumer education classes and will see and hear more consumer education messages. This new socialization process can be expected to produce consumers who will be more willing and able to state what they want and need to sellers and who will also be more willing and able to file complaints and lawsuits against sellers. In addition, this process can be expected to produce consumers who will seek out and use more information before making purchases and who will tend to make an increasing amount of these purchases from sellers who (1) provide large amounts of easily-acquired consumer information, (2) conduct consumer education programs, and (3) offer products that present little danger to the public's health.

Finally, this process can be expected to produce consumers who will fight more actively in the political arena for legislation which serves their interests. All of these hypotheses should be tested by conducting cross-sectional and longitudinal studies of how the many present-day consumer education programs have affected consumer behavior.

#### References

- Brooks, Thomas M. "Consumer Education: Can We Improve Our Score?" Journal of Home Economics, 65 (September, 1973), 33-35.
- Day, George S. and William K. Brandt. "Consumer Research and the Evaluation of Information Disclosure Requirements: The Case of Truth in Lending," Journal of Consumer Research, 1 (June, 1974), 21-32.
- "Divorce Course," Time (December 2, 1974), p. 92.
- Goetting, Marsha. "Let's Put Consumer Education in Prison," Journal of Consumer Affairs, 8 (Winter, 1974), 198-203.
- Isakson, Hans R. and Alex R. Maurizi. "The Consumer Economics of Unit Pricing," Journal of Marketing Research 10 (August, 1973), 277-85.
- Jacoby, Jacob, Donald E. Speller, and Carol A. Kohn. "Brand Choice Behavior As A Function of Information Load," Journal of Marketing Research, 11 (February, 1974), 63-9.
- Johnson, Edna D. "Remarks by Edna D. Johnson," A speech given before the White House conference on consumer education, March 11, 1975.

Levy, Claudia. "Giant's Net Shows Gain Over '74," Washington Post (June 4, 1974), p. D1.

Mahan, Linda. "Twelfth-Year Itch," Consumer Education Forum, 3 (Spring, 1972), 1.

Office of Consumer Affairs. "Directory of Consumer Education Activities Conducted by State, County, and City Consumer Agencies," mimeo, U.S. Dept. of Health, Education, and Welfare, June, 1974.

\_\_\_\_\_. "Federally-Funded Consumer Programs in Low-Income Communities: A Preliminary Compilation," mimeo, U.S. Dept. of Health, Education, and Welfare, May, 1973.

Peterson, Esther. "Consumerism As a Retailer's Asset," Harvard Business Review, 52 (May-June, 1974), 91-101.

Scherf, Gerhard, W.H. "Consumer Education As A Means of Alleviating Dissatisfaction," Journal of Consumer Affairs, 8 (Summer, 1974), 61-75.

Seitz, Wesley D. "Consumer Education As the Means to Attain Efficient Market Performance," Journal of Consumer Affairs, 6 (Winter, 1972), 199-201.

Uhl, Joseph N. and Others. Survey and Evaluation of Consumer Education Programs in the United States.

Willett, Sandra L. "Preliminary Statement," A handout prepared for the Panel on Consumer Education of the 1974 meeting of the Education Commission of the States, June 21, 1974.

## REFLECTIONS ON RESEARCH IN CONSUMER BEHAVIOR

Johan Arndt, Norwegian School of Economics and Business Administration

### Abstract

This review of the published research in consumer behavior concludes that the area has been incompletely covered by the research effort. Most of the research attention has been given to prepurchase decision processes for brands. More fundamental consumer problems such as purchase of strategic products and budget allocation problems, as well as postpurchase phenomena in general, have been neglected. It is suggested that the bias in the research attention may be traced back to the dominance of a pragmatic marketing tradition in the area, encouraging emphasis on controllable aspects of the decision process.

### Introduction

The purpose of this paper is to make some general comments on selected aspects of consumer behavior research. Specifically, an attempt to be made to argue for more emphasis on theory based research on what is called "fundamental consumer problems".

The paper which builds on the draft of Chapter One in a forthcoming book tentatively titled Toward a Theory of Household Consumer Behavior, is organized as follows:

- The area of consumer behavior.
- Parties interested in consumer research.
- Critical comments on the consumer behavior literature.
- The task ahead.

### The Area of Consumer Behavior

Before proceeding to the discussion of the subject matter, it is necessary to state as clearly as possible what is meant by consumer behavior. In wider sense, consumer behavior problems are the problems encountered by members of society in the acquisition and realization of their standard of living. More specifically, consumer behavior may be defined as (1) the mental and physical acts of (2) individuals or organized groups of individuals concerned with ultimate consumption, (3) that relate to acquiring, using, and, in some cases, disposing of (4) economic goods and services (5) both from the private and the public sector.

In order to delineate the area of consumer behavior further, the various activities constituting "consumer behavior" may be cross-classified by (1) what may be called stage in the buy-use continuum and (2) by level of process.

Many writers, such as Nicosia (1966), Engel, Kollat, and Blackwell (1968, 1973), Howard and Sheth (1969), Robertson (1971), and Hansen (1972) have viewed consumer decision making as a multistage, problem-solving process. Building on Engel, Kollat, and Blackwell (1968, 1973), the following stages may be distinguished.<sup>1</sup>

<sup>1</sup>It should be noted that Engel, Kollat, and Blackwell (1968, 1973) are mainly interested in the purchasing process. Hence, they do not include steps 4 and 5 in the model below. Instead, the purchasing stage is followed by postpurchase evaluation and cognitive dissonance following in the wake of the decision.

- (1) Problem-recognition.
- (2) Search for information to evaluate alternatives.
  - (a) Internal search.
  - (b) External search.
- (3) Implementation of the purchase.
- (4) Physical consumption.
- (5) Postconsumption activities.

At the problem-recognition stage, the consumer becomes aware of a difference between the desired state with regard to the fulfillment of some consumer goal and the actual state. It seems reasonable to assume that this difference has to reach a certain threshold value in order to warrant further action.

If the discrepancy between the desired and the actual state is of a sufficient magnitude, the consumer will move into the second stage where information is sought to locate purchase alternatives, to clarify buying goals (evaluation criteria), and to determine to what extent the alternatives measure up to the goals. In this phase, the consumer may start by determining whether there is sufficient stored information that may be generalized to the problem at hand. If not, and the problem is important enough, the consumer may start an external search process, for instance by inspecting products directly (significant information) and/or by acquiring symbolic information representing the product by means of commercial mass media, word-of-mouth communications from other consumers, or information from neutral sources such as Consumer Reports.

In the implementation stage, the transaction process by which the right of property is transferred, is completed.

The fourth stage, physical consumption, is concerned with the use of the product or service to satisfy the buying goals. For non-durable goods and services, the consumption may be completed in a short time. For durables, such as appliances and houses, the consumption may occur continuously over many time periods. In practice, the actual physical depreciation in a given period may be hard to determine.

Finally, during or after actual consumption, the consumer may compare the degree of goal-fulfillment with the original expectations. The degree of satisfaction felt is then believed to be stored in the consumer's mind as part of an "attitude". This phase also contains attempts to cope with postpurchase regrets or cognitive dissonance, as well as activities to remove remainders of the goods consumed from the household, for instance in the form of garbage or a used car traded in.

The five-stage process should not be conceived of as being necessarily unidirectional in a simple fashion. First, it is reasonable to believe that there are feedback relationships between "later" and "earlier" stages. For example, postconsumption activities may affect problem-recognition processes in the next cycle. Second, the actual sequence of stages may deviate from the regular pattern. For instance, stage 3 may technically occur before stage 2b if a product is bought on trial as a first part of a search process.

The second dimension in our scheme, level of decision process, relates to the notion that certain decisions

depend on other decisions. Important key purchasing decisions are believed to create constraints or criteria for subsequent minor decisions. Hence, decisions may be placed in a hierarchy on basis of the extent to which the decisions depend on other decisions. The levels suggested below build to some extent on the work of the Danish economist Karen Gredal (1964, 1966):

- A. Acquisition of strategic items.
- B. Central budget allocation decisions.
- C. Generic product or service decisions.
- D. Variant selection.<sup>2</sup>

Certain buying decisions are strategic in the sense that they are concerned with long-term bindings of resources. Hereby, they affect the budget available for other goods and services. Examples are expensive durable goods such as houses, cars, etc. Second, a decision may be considered strategic since purchase of one item usually entails purchase of other items. For instance, a well-to-do family may have a choice between sending their son or daughter to a prestigious private college, acquiring a cottage in a mountain resort, or moving to a more attractive neighborhood. In this case, choice of alternative would have important repercussions for other items to be purchased. For instance, the cottage alternative may result in acquisition of hiking equipment, fishing tackle, etc.

The central budget allocation decisions correspond to what Gredal (1964, 1966) terms "the general purchasing decision". The purpose of decisions at this level is to establish more or less explicit policies about the allocation of the available money among alternative groups of products and services. It is possible that, at this level, appropriation is first made for fixed expenses such as interest and installments for debts, costs of utilities, cost of maintenance of house, etc.

Level C, which parallels Gredal's (1964, 1966) notion of "the concrete purchasing decision", is concerned, as the term suggests, with selection of generic product and service categories and not, for instance, with the individual brands.

Selection of alternatives within each product or service group occurs at the fourth level, which in Gredal's (1964, 1966) terms is called "the selection decision".

The four-level hierarchy presented here should not be thought of as a procedure that all consumers necessarily follow. It is conceivable that certain levels may be skipped. For instance, the decision maker may move directly from level B to level D, without making any generic product or service decision. However, the normal rule is that decisions at the higher levels in the hierarchy occur before and occur less frequently than decisions at the lower levels. Hence, our typology parallels the distinction between strategic decisions and operational decisions in the business management literature.

Figure 1 presents a dummy matrix showing the various cells when stages and levels in our system of consumer behavior are cross-classified. While all cells should not be considered being of equal importance as regards research, it is suggested that the research effort so

far has been very lop-sided. As indicated in Figure 1, it seems as if most of the published consumer research material has focused on cells D-1, D-2, and D-3. An attempt will be made to give a systematic explanation of this bias.

#### Parties Interested in Consumer Research

It is almost axiomatic that the research activities in any academic area do not occur in what may be called an "influence vacuum". The direction of the consumer research is believed not only to be influenced by consumer researchers themselves, but also by the interests and influences of other groups and institutions in society, who are current or potential users of consumer behavior knowledge.

This section will point out some of these interest groups of relevance for consumer research and tentatively suggest the nature of their interest in the subject matter. In Figure 2 below, the main goals and objectives for the most important groups are suggested. Second, building on the thoughts of Sommers (1971), the kind of validity searched for in the research questions raised is also suggested.

As pointed out by Sommers (1971), and as indicated in Figure 2, marketing managers tend to be governed by a prescriptive tradition, that is, determining the firm's strategy for a given product or brand. As the manager is normally rewarded on the basis of the economic consequences of his manipulations of the variables in the marketing mix, he is normally not interested in the research available per se, but in the extent to which knowledge about consumer behavior may be converted into profit-making practice. Hence, he is oriented toward the pragmatic validity of the relationships uncovered.

Similarly, commercial marketing researchers are often atheoretical in the approach taken as they perceive their role as providing managers with information bearing directly on the decisions to be made. Though researchers in many ways share other practitioners' pragmatic interest, they usually attempt to widen the scope of managers by moving further to more detailed descriptions of market behavior of consumers. Hereby, as stated by Sommers (1971), commercial marketing researchers are concerned with concurrent validity - making estimates of the magnitude of present behavior - and content validity - determining the fairness and accuracy of the measurements of the behavior of interest.

As regards academic consumer behavior researchers, Sommers (1971) claims that this type is concerned with questions like, "What is it that is working and how?", and that this interest in explanation of aspects of consumer behavior is tantamount to a concern with construct validity. However, the pragmatic orientation of marketers seems also to pervade this party in the sense that, so far, there has been little emphasis on theory building. The research effort has been biased because of an overemphasis on prepurchase variables. Therefore, the consumer researchers have so far been too influenced by the marketing management thinking and have not really reached the construct validity level.

To date, academic behavioral researchers in areas other than marketing have not played a major part in the development of the field. This party is very heterogeneous, and it is difficult to make meaningful generalizations. However, there appear to be two main differing schools of thought. First, there are "the traditionalists" interested in explaining, predicting and theorizing about behavior - in other words, in construct validity. There is also a so-called "radical"

<sup>2</sup>Gredal (1964, 1966) does not include level A - Acquisition of strategic items. On the other hand she has added a final level termed "the technical purchasing action". In this scheme, however, this level seems to correspond to the implementation stage in the stage dimension.

FIGURE 1  
ELEMENTS OF CONSUMER BEHAVIOR

Level	Stage	1	2	3	4	5
		Problem- Recogni- tion	Search for information Internal a	External b	Implementa- tion of the purchase	Physical consump- tion
A. Acquisition of strategic items						
B. Central budget allocation decisions						
C. Generic product or service decisions						
D. Variant selection		x	x	x	x	

FIGURE 2  
PARTIES INTERESTED IN CONSUMER BEHAVIOR, THEIR MAIN GOALS, AND TYPE OF VALIDITY DESIRED

Party	Main goal	Type of validity desired
Marketing managers	To control consumer brand choice by influencing decision processes	Pragmatic validity
Commercial marketing researchers	To describe controllable variables in consumer decision processes. To estimate quantities of actual or potential sales	Concurrent and content validity
Academic consumer behavior researchers in marketing	To identify and describe variables of interest for marketing managers	Concurrent and content validity
Academic behavioral scientists	Goals depend on research tradition adhered to	
	Traditional orientation: to explain consumer behavior	Construct validity
Authorities	Radical orientation: to emancipate consumers	Pragmatic validity
	To describe consumer interests in order to allocate resources among different public expenditure alternatives and to formulate regulatory measures	Pragmatic validity
Consumer educators	To help consumers make more rational decisions	Pragmatic validity
Consumers	To be able to make better purchasing decisions	Pragmatic validity

school which is more interested in action than in description. Specifically, this tradition may be said to be oriented toward the emancipation of consumers from being objects manipulated by profit-seeking producers. In this sense, this party is interested in pragmatic validity, but from a standpoint opposite to that of marketing managers.

The primary concern of the authorities in most countries is likewise with pragmatic action rather than with theory.

Consumer educators are in many ways similar to the

"radical" academic behavioral scientists in the sense that their goal is to help consumers. However, in contrast to the latter group, this group is normally in immediate grass roots contact with consumers and work mainly within the present power structure.

Finally, there is the heterogeneous group of consumers who so far may have had little direct benefit from the research done. The research has been governed by consumer interests only to a small extent, and research reports have seldom been prepared for consumer audiences. In most cases, the target groups for research products have been other academicians or business

firms. The object of this research has been neglected as a target group. One of the few marketing writers discussing how consumers might use consumer behavior knowledge is Nicosia (1969).

#### Critical Comments on the Consumer Behavior Literature

A perusal of the literature in consumer behavior leaves little doubt that marketing has been the dominant contributing field. Only in marketing there has been anything approaching a research tradition. The few interdisciplinary ventures launched in consumer behavior have to a great extent been initiated and dominated by marketing academicians. For instance, the Association for Consumer Research founded in 1970, as well as the Journal of Consumer Research introduced in 1974, seem to support the marketing perspective of consumer behavior by several criteria.<sup>3</sup>

Therefore, the following comments will pertain particularly to marketing publications in consumer behavior. The comments are deliberately biased in the sense that they accentuate the shortcomings of the literature rather than the advances made.

In spite of the quantitatively impressive literature tonnage, the marketing contributions to consumer behavior seem to be flawed by six serious shortcomings:

- Neglect of consumer behavior with regard to the public sector.
- Fragmentation of the subject matter.
- Inadequate attention to consumer behavior of groups.
- Bias toward decision processes.
- Insignificant problems.
- Irrelevance of findings.

#### Neglect of Consumer Behavior with Regard to the Public Sector

This review of the consumer behavior literature shows that the research effort, with a negligible number of exceptions, has focused on the encounter of the consumer with the offerings of private sector of the economy. This means that the consumer behavior discipline has "defined away" the consumer problems relating to the functioning of important public services such as transportation, education, and health.

A possible reason for this self-imposed limit on the subject matter may be the general neglect of the public sector in most Western countries, a point raised provocatively by Galbraith (1958). However, a more plausible reason may be the marketing dominance of the consumer behavior field earlier pointed out. Since the public sector normally is not regulated by market mechanisms, marketing researchers may not have recognized "supply-demand" relations for public services as marketing problems.

One noteworthy attempt to broaden the scope of marketing and consumer behavior is the contribution of Kotler and Zaltman (1971) in applying marketing techniques and concepts such as "product", "promotion", and "price" to nontraditional areas such as family planning, fund-raising for medical research, and political campaigns.

<sup>3</sup>For example, the editor plus about one-half of the members of the editorial board of the Journal of Consumer Research may apparently be classified as marketing men. Moreover, at least 6 out of the 9 articles in the first issue of the journal seem to be contributions in the marketing tradition.

However, at its present stage of development, this "broadening" is more related to new uses of existing marketing techniques and philosophies rather than to the development of further knowledge of consumer behavior.

#### Fragmentation of the Subject Matter

Judged by the volume of output, the area of consumer behavior is healthy indeed. However, productivity is not the same as real progress. The great majority of the studies may be classified as non-cumulative empirical-analytical research investigations at the micro level. Most research studies have used what Nicosia (1966, pp. 11-12) calls "reduced form" models or selected constructs isolated from consumer behavior as a whole. As pointed out by Kollat, Engel, and Blackwell (1970, p. 328), consumer research projects tend to be "the result of the availability of data, the convenience of research and mathematical techniques, and/or the appeal of certain behavioral constructs. In other words, most research has been data-technique-construct motivated and oriented and has typically been conducted independently with little apparent coordination."

Such opportunism and reductionism have subordinated conceptualization and theory-building to measurement and manipulation of data by high-powered statistical tools. This has resulted in a large number of isolated facts, which lack consistency and which are difficult to integrate into formal comprehensive theories. Therefore, it is not difficult to agree with Sheth (1967, p. 742), who concluding his review of the buyer behavior literature, used the analogy of the seven blind men touching different parts of the elephant to characterize the level of theoretical development in the area.

Only a few of the books of consumer behavior contains attempts to formulate comprehensive theories integrating the field, see Nicosia (1966), Engel, Kollat, and Blackwell (1968, 1973), Howard and Sheth (1969), and Hansen (1972). However, there are publications dealing with selected aspects of consumer behavior such as social class, see Carman (1965); communications behavior, see Nowak, Carlman, and Wärneryd (1966); brand image and attitudes, see Myers (1968); and new product diffusion, see Robertson (1971). There are also a few comprehensive reviews of the literature such as Howard (1963) and Burk (1968) and the deep-plowing readings books published by Ward and Robertson (1973) and Sheth (1974) of original papers in consumer behavior research.

If we disregard textbooks such as Myers and Reynolds (1967), Markin (1969, 1974), Sandell (1969), Bliss (1970, Robertson (1970), Walters and Paul (1970), Bennett and Kassarian (1972), and Walters (1974), the rest of the books uncovered are readings books, presenting papers from general symposia of consumer behavior; see Clark (1955a, 1958), Newman (1966), Arndt (1968), Sommers and Kernan (1968), Gardner (1971), Venkatesan (1972), Ward and Wright (1974), and Schlinger (1975); symposia on various aspects of consumer behavior, see Clark (1955b), Foote (1961), Adler and Crespi (1966, 1968), King and Tigert (1970), and Haley (1972); or organized research programs, such as Cox (1967) and Farley, Howard, and Ring (1974). The remainder is simply more or less logically organized readings books of papers already published. Examples are Bliss (1963, 1967), Grossack (1964), McNeal (1965), Britt (1966, 1970a, 1970b), Kassarian and Robertson (1968, 1973), Engel (1968), Kollat, Blackwell, and Engel (1970), Aaker and Day (1971), Day and Ness (1971), Holloway, Mittelstaedt, and Venkatesan (1971), Cohen (1972), Grønhaug (1972), and Howard and Ostlund (1973). While

such proliferation of readings may have provided a temporary relief to authors threatened by the "publish or perish" syndrome, it has added little to the structuring of the field.

In conclusion, there is little doubt that the fragmentation of the subject matter is a symptom of the absence of adequate theoretical underpinnings in the area which could give direction and meaning to empirical research. Hence, in reiterating the need for comprehensive, integrative theory-building, the present writer joins reviewers such as Perloff (1968), Kollat, Engel, and Balckwell (1970), Kollat, Blackwell, and Engel (1968), Pollay (1972), and Robertson and Ward (1973a, 1973b), as well as apparently even the majority of the membership of the Association for Consumer Research, see Ford, Kuehl, and Dyer (1975).

#### Inadequate Attention to Consumer Behavior of Groups

The few existing attempts to formulate integrative models of consumer behavior - Nicosia (1966), Engel, Kollat, and Balckwell (1968, 1973), Howard and Sheth (1969), and Hansen (1972) - have approached consumer problems at the individual, rather than at the group level. With the exception of discussion in Hansen (1972, pp. 409-431), the problems of moving from behavior at the individual level to having groups of consumers as the unit of study have been more or less ignored. However, both from the studies in intra-familial decision-making, see Davis and Rigaux (1974), and the diffusion of innovation studies, see Rogers and Shoemaker (1971), it seems clear that it is theoretically incorrect to model group processes simply by adding individual behavior. While in the future the behavior of individuals is still expected to be of interest to the consumer behavior theorist, theories should also be constructed at the group level.

#### Bias toward Decision Processes

The area is also flawed by the shortcoming of overemphasis on decision processes and what may be called "short-term effects of marketing mix" studies. As was concluded above, this means neglect of what happens after purchase including the consequences of consumption, both for the decision unit itself and for the greater society. For instance, the possibility that obesity may activate diabetes, should not be considered a problem of medicine only. To the extent that this is a problem caused by certain patterns of consumption, it should also be of interest to consumer behaviorists. Similarly, a neglect of consequences occurs at the macro level. For instance, Feldman (1971, p. 55) charged the marketing discipline of failure to recognize that (the) "products, which are marketing outputs designed for individual satisfaction, are simultaneously inputs to a larger environmental system and as such may affect the well-being of society".

#### Insignificant Problems

Another bias in the research is the preoccupation with decisions that would appear to be of minor interest to the consumer, such as brand choice for convenience goods, and the corresponding neglect of more fundamental decision areas. While research attention has been given to issues like the proverbial 1-ply vs. 2-ply toilet tissue decision, more important areas such as the emergence of new value systems and their consequences, and problems relating to nutrition of disadvantaged consumer groups have to a great extent been ignored.

#### Irrelevance of Findings

This myopia and exaggerated pragmatism have led to what

seems to be a sixth main shortcoming - irrelevance of the findings. The production of isolated facts and relationships by testing narrow hypotheses not derived by theoretical considerations seems to have resulted in a data base which not only is trivial, but also is irrelevant, and hence is not much applicable as basis for managerial action.<sup>4</sup>

#### The Image of Man in Consumer Research

In conclusion, the image of man as a consumer emerging from the consumer behavior literature is in many ways reminiscent of Kover's (1967) model of man as defined by commercial marketing research:

Man is alone, an isolate.

Man is passive.

Man is mainly concerned with publicly acceptable needs and appearance.

Men's reactions can be explained or classified by their styles of life - by earnings, age, sex and education.

Products and services are important in meeting the needs of man.

People can adequately express their feelings about advertising and products.

Although the consumer can talk clearly and at length about products, he does not know what these beliefs mean and which beliefs affect him most.

Moreover, to this might be added:

Man is concerned with present rather than with future needs.

Man is mostly interested in minor decisions such as brand selection. More basic decisions are of minor importance.

It is tempting to conclude that this state of the art may to a large extent be due to the domination of the marketing perspective in the current thinking on consumer behavior. Little imagination is needed to see the similarity between the one-dimensional consumer depicted above and the docile super-consuming being believed to the ideal customer from the producers' viewpoint. Hence, normative and positive models of consumers may have converged.

#### The Task Ahead

By several criteria, the consumer behavior area gives the impression of having become a science in its own right. The area has a rapidly growing professional organization, at least one specialized journal, and an exponentially increasing body of literature, and enjoys the academic autonomy of being represented by specialized university courses. Several writers, among them Nicosia (1969), Kollat, Blackwell, and Engel (1972),

<sup>4</sup>A good illustration of this aspect of the state of the art is found in Tigert's (1972) scathing critique of one of the winners of the American Marketing Association's 1969-70 Doctoral Dissertation Competition. While crediting the researcher for competence in statistics and knowledge of research methods, Tigert claims that neither he nor the author knows what to do with the results of the analysis - clusters of consumers extracted by cluster analysis. Tigert concludes in no uncertain terms, "That his talents have been wasted on a fruitless exercise is scandalous, a total disaster ---Gentlemen, the name of the game is relevance. How can we expect the business community to support doctoral research if we continue to work on trivial problems?"



and Sheth (1974), have noted the impressive gains in systematic knowledge in the area. In the words of Tucker (1968, p. 267), consumer behavior has broken away from pseudo-discipline. Sheth (1974) is even so optimistic as to suggest that within a decade, the area may begin to reverse the flow of borrowing by "exporting" consumer behavior theory both to the more mature and to the less mature social sciences.

However, an evaluation of the state of the art of an area must deal primarily with its present status, and not with its humble beginnings or with its conceivable future potentials. And in this respect, it has been concluded that consumer behavior has yet to establish the theoretical cohesiveness and the unique research tradition indicative of a mature discipline. The very fact that the question of whether or not consumer behavior is a science or a unique discipline is constantly and self-consciously asked, signals lack of self-confidence and professional adolescence.

Commenting on the academic autonomy of the area of consumer behavior, writers such as Tucker (1968) and Kas-sarjian and Robertson (1968, p. 2) define consumer behavior as a part of the marketing decision-making system. Sheth (1974, p. 395), on the other hand, views consumer behavior as a unique area, which is a subsystem neither of marketing nor of any of the existing social sciences. In the opinion of the present writer, the claims of uniqueness for the consumer behavior area are hardly substantiated, in view of the apparent dependence on the marketing discipline for the perspectives and goals of the area, and the almost complete indebtedness to the behavior sciences for conceptual and methodological tools.

Therefore, as pointed out by Pollay (1972, p. 595), there is so far too much evidence of disciplinary immaturity in consumer behavior: proliferation of theories or partial theories, expressions of disagreement over the priority of problems, a faster growth in complexity than accuracy, and professional insecurity among researchers in the area. In the terms of Kuhn (1962), consumer behavior is still a field of study in a pre-paradigm state.

Symptomatic of the present situation is the identity crisis of many consumer behaviorists confronted with what Robertson and Ward (1973b, p. 60) call the pay-off dilemma - the fact that consumer researchers seek to engage in theoretical research and at the same time to meet the need for immediately actionable results of dominant user groups such as marketing practitioners. It follows from the criticism in this paper that few consumer behaviorists have handled this role conflict satisfactorily. Hence, it is no wonder that the nagging question of relevance keeps popping up: Are the results produced by current consumer research relevant to the problems faced by marketers as well as consumers, both at the micro and macro level?

To advance the state of the art, several suggestions have been made. Pollay (1972, p. 595) has suggested that a possible reason for the state of consumer research is the fact that the area may not have attracted the top caliber scholars. Though it is hard to quarrel with the contention that, with a few exceptions, the consumer behavior area gives the impression of being mainly occupied by "little thinkers", it is still somewhat fatalistic to wait passively for a Messiah to appear.

One possibility suggesting itself is to continue the production of research results in the present "anarchical" fashion by gradually sharpening the tools of measurement and analysis and hoping that knowledge will

develop in a cumulative fashion. However, in the development of sciences, development-by-revolution is often more important than development-by-accumulation, at least in the natural science areas, see Kuhn (1962). Therefore, research productivity should not be mistaken for genuine progress. There is little basis for any hope that the sheer persistence in the data gathering may eventually lead to conceptual break-throughs. Nor is it realistic to expect that in the "market-place of research", the "good research" will eventually drive out the "bad research". So far in consumer research, Gresham's law has apparently easily neutralized the law of Darwin. However, not deviating much from the present research tradition, it is possible to advance consumer research by following the advice of Glock and Nicosia (1963, pp. 25-26) who advocated more attention to studies of strategic populations, that is observations of consumers as they are in the midst of making not one choice, but a wide variety of choices. Examples of this approach are studies of consumers who have just moved to a new area (Andreassen, 1966; and Bell, 1969), consumer socialization studies (Ward, 1974), and studies of total leisure time activities, see the papers in the 1974 Conference of the Association for Consumer Research (Schlinger, 1975).

Kollat, Engel, and Blackwell (1970), and Kollat, Blackwell, and Engel (1972) propose establishing research priorities, that is identifying what aspects of consumer behavior that are of the greatest importance and what phenomena need to be investigated so that these key areas can be understood. Similarly, Pollay (1972) suggests organized research programs. Though it is hard to quarrel with the logic of this position, it is clear that this solution is less than complete as long as critical questions remain unanswered: Who are to determine the research priorities? What criteria should be used to evaluate the potential importance of an aspect of consumer behavior? And who should be the reference group when deciding on importance - marketers, consumers, consumerist advocates, or government? Another caveat against premature programming is that a uniformed effort may magnify consequences of mistakes made.

Several writers including Sheth (1967), Pollay (1972), and Zaltman, Pinson, and Angelmar (1973) have made a convincing case for more emphasis on metatheoretic aspects of theory building, that is for development of a rigorous thinking methodology to evaluate theories. Yet, while the present achievements and future potentials of methatheory are not to be denied, the main problems as of now in consumer behavior are problems of content and substance of theory rather than its form.

Past theories give the impression of having been constructed by merging empirical findings from narrow research studies with borrowed behavioral science findings and concepts. This build-up approach to theory construction together with the marketing frame of reference dominating the area, may partly be responsible for generating the instrumental, passive model of man depicted in the last section.

As soon as possible, a challenge should be made to the dominating positivistic research philosophy, with its main emphasis on data gathering and data analysis. What is above all needed is a new vision of consumer behavior - a specification of the domain of consumer behavior broadening the field of study and revising the current explicit and implicit research priorities. The "self-evident" orientation toward marketing practitioners as the key reference group could well be replaced by an endorsement of a true consumer frame of reference. To follow the advice of Perloff (Twedt, 1965, p. 266) for consumer psychologists, consumer research-

ers should seek directly and explicitly to serve the consumer's needs, to study the consumer as a consumer rather than as an individual whose attention and purchasing behavior are coveted to maximize profits. This means that research priorities should be governed by the relative importance of the unsolved problems for the consumer. Some clues to the relative importance may be found in the conceptual framework implicit in the level of decision process/stage in the buy-use continuum presented above in this paper. If the monetary and social importance of the decision to the consumer is to be the main criterion, this would mean more emphasis on what was called acquisition of strategic items and central budget allocation decisions. At the macro level, socially important issues such as the social and ecological consequences of consumption structures should also be defined as legitimate areas for study of the consumer behaviorists.

It may be that a chief cause of the paucity or absence of theoretical underpinnings in consumer behavior is its status as having mainly been a part of marketing. As long as consumer behavior remains an appendage of marketing, it will be under pressure to remain an applied field, concerned with the delivery of existing knowledge to solve marketing problems. Since an applied orientation almost by necessity must be in conflict with aspirations of building a solid theoretical base, perhaps consumer behavior may only become of discipline in its own right by cutting its umbilical cord to marketing. Such an emancipation may contribute to solve the pay-off dilemma of consumer behavior researchers by removing the pressure to produce managerially actionable results, and hence to allow an all-out effort on conceptualization. Analysis of buyers from marketing's viewpoint should remain within marketing and could be given the more appropriate term "customer analysis".

In the short run, the results of such a broadened, theory-based consumer behavior discipline may appear to fail the immediate action needs of marketers. In the long run, however, it is possible that the insights gained by more fundamental theories may well turn out to be even more actionable than the current narrower models.

#### References

- Aaker, David A. and George S. Day (Eds.) Consumerism: Search for the Consumer Interest. New York: The Free Press, 1971.
- Adler, Lee and Irving Crespi (Eds.) Attitude Research at Sea. Chicago: American Marketing Association, 1966.
- Adler, Lee and Irving Crespi (Eds.) Attitude Research on the Rocks. Chicago: American Marketing Association, 1968.
- Andreasen, Alan R. "Geographic Mobility and Market Segmentation," Journal of Marketing Research, 3 (November 1966), 341-8.
- Arndt, Johan (Ed.) Insights into Consumer Behavior. Boston: Allyn and Bacon, 1968.
- Bell, James E., Jr. "Mobiles - a Neglected Market Segment," Journal of Marketing, 33 (April 1969), 37-44.
- Bennett, Peter D. and Harold H. Kassarian. Consumer Behavior. Englewood Cliffs: Prentice-Hall, 1972.
- Bliss, Perry (Ed.) Marketing and the Behavioral Sciences. Boston: Allyn and Bacon, 1963.
- Bliss, Perry (Ed.) Marketing and the Behavioral Sciences. 2nd Edition. Boston: Allyn and Bacon, 1967.
- Bliss, Perry. Marketing Management and the Behavioral Environment. Englewood Cliffs: Prentice-Hall, 1970.
- Britt, Steuart H. (Ed.) Consumer Behavior and the Behavioral Sciences. New York: John Wiley and Sons, 1966.
- Britt, Steuart H. (Ed.) Consumer Behavior in Theory and in Action. New York: John Wiley and Sons, 1970a.
- Britt, Steuart H. (Ed.) Psychological Experiments in Consumer Behavior. New York: John Wiley and Sons, 1970b.
- Burk, Marguerite C. Consumption Economics: A Multidisciplinary Approach. New York: John Wiley and Sons, 1968.
- Carman, James M. The Application of Social Class in Market Segmentation. Berkeley: Institute of Business and Economic Research, Graduate School of Business Administration, University of California, 1965.
- Clark, Lincoln H. (Ed.) Consumer Behavior: The Dynamics of Consumer Reaction. New York: New York University Press, 1955a.
- Clark, Lincoln H. (Ed.) Consumer Behavior, Vol. II. The Life Cycle and Consumer Behavior. New York: New York University Press, 1955b.
- Clark, Lincoln H. (Ed.) Consumer Behavior: Research on Consumer Reactions. New York: Harper & Brothers, 1958.
- Cohen, Joel B. (Ed.) Behavioral Science Foundations of Consumer Behavior. New York: The Free Press, 1972.
- Cox, Donald F. (Ed.) Risk Taking and Information Handling in Consumer Behavior. Boston: Division of Research, Graduate School of Business Administration, Harvard University, 1967.
- Davis, Harry L. and Benny P. Rigaux. "Perception of Marital Roles in Decision Processes," Journal of Consumer Research, 1 (June 1974), 51-62.
- Day, Ralph L. and Thomas E. Ness (Eds.) Marketing Models: Behavioral Science Applications. Scranton: International Textbook Company, 1971.
- Engel, James F. (Ed.) Consumer Behavior: Selected Readings. Homewood: Richard D. Irwin, 1968.
- Engel, James F., David T. Kollat, and Roger D. Blackwell. Consumer Behavior. New York: Holt, Rinehart and Winston, 1968.
- Engel, James F., David T. Kollat, and Roger D. Blackwell. Consumer Behavior. Second Edition. New York: Holt, Rinehart and Winston, 1973.
- Farley, John U., John A. Howard, and L. Winston Ring (Eds.) Consumer Behavior: Theory and Application. Boston: Allyn and Bacon, 1974.
- Feldman, Laurence P. "Societal Adaptation: A New Challenge for Marketing," Journal of Marketing, 35 (July 1971), 54-60.
- Foote, Nelson N. (Ed.) Household Decision-Making. Consumer Behavior, Vol IV. New York: New York University Press, 1961.

- Ford, Gary T., Philip G. Kuehl, and Robert P. Dyer. "The Status of Consumer Behavior: Some Empirical Perspectives," in Mary J. Schlinger (Ed.), Advances in Consumer Research, Vol. 2. Chicago: Association for Consumer Research, 1975, 51-61.
- Galbraith, John K. The Affluent Society. Boston: Houghton-Mifflin Company, 1958.
- Gardner, David M. (Ed.) Proceedings. 2nd Annual Conference. College Park, Maryland: Association for Consumer Research, 1971.
- Glock, Charles Y. and Francesco M. Nicosia. "Sociology and the Study of Consumers," Journal of Advertising Research, 3 (September 1963), 21-7.
- Gredal, Karen. Moderne forbrugeres motiver og adferd. København: Einar Harcks Forlag, 1964.
- Gredal, Karen. "Purchasing Behaviour in Households," in Max Kjaer-Hansen (Ed.), Readings in Danish Theory of Marketing. København: Einar Harcks Forlag, 1966, 84-100.
- Grønhaug, Kjell (Ed.) Forbrukeratferd. Oslo: Johan Grundt Tanum, 1972.
- Grossack, Martin M. (Ed.) Understanding Consumer Behavior. Boston: The Christopher Publishing House, 1964.
- Haley, Russell I. (Ed.) Attitude Research in Transition. Chicago: American Marketing Association, 1972.
- Hansen, Flemming. Consumer Choice Behavior: A Cognitive Theory. New York: The Free Press, 1972.
- Holloway, Robert J., Robert A. Mittelstaedt, and M. Venkatesan (Eds.) Consumer Behavior: Contemporary Research in Action. Boston: Houghton Mifflin Company, 1971.
- Howard, John A. Marketing: Executive and Buyer Behavior. New York: Columbia University Press, 1963.
- Howard, John A. and Lyman E. Ostlund (Eds.) Buyer Behavior. New York: Alfred A. Knopf, 1973.
- Howard, John A. and Jagdish N. Sheth. The Theory of Buyer Behavior. New York: John Wiley and Sons, 1969.
- Kassarjian, Harold H. and Thomas S. Robertson (Eds.) Perspectives in Consumer Behavior. Glenview: Scott, Foresman and Company, 1968.
- Kassarjian, Harold H. and Thomas S. Robertson (Eds.) Perspectives in Consumer Behavior. Revised Edition. Glenview: Scott, Foresman and Company, 1973.
- King, Charles W. and Douglas J. Tigert (Eds.) Attitude Research Reaches New Heights. Chicago: American Marketing Association, 1970.
- Kollat, David T., Roger D. Blackwell, and James F. Engel (Eds.) Research in Consumer Behavior. New York: Holt, Rinehart and Winston, 1970.
- Kollat, David T., Roger D. Blackwell, and James F. Engel. "The Current Status of Consumer Behavior Research: Developments during the 1968-1972 Period," in M. Venkatesan (Ed.), Proceedings. 3rd Annual Conference. Iowa City: Association for Consumer Research, 1972, 576-85.
- Kollat, David T., James F. Engel, and Roger D. Blackwell. "Current Problems in Consumer Behavior Research," Journal of Marketing Research, 2 (August 1970), 327-32.
- Kotler, Philip and Gerald Zaltman. "Social Marketing: An Approach to Planned Social Change," Journal of Marketing, 35 (July 1971), 3-12.
- Kover, Arthur J. "Models of Man as Defined by Marketing Research," Journal of Marketing Research, 4 (May 1967), 129-32.
- Kuhn, Thomas S. The Structure of Scientific Revolutions. Chicago: University of Chicago Press, 1962.
- Markin, Rom J. The Psychology of Consumer Behavior. Englewood Cliffs: Prentice-Hall, 1969.
- Markin, Rom J., Jr. Consumer Behavior: A Cognitive Orientation. New York: MacMillan Publishing Company, 1974.
- McNeal, James U. (Ed.) Dimensions of Consumer Behavior. New York: Appleton-Century-Crofts, 1965.
- Myers, James H. and William H. Reynolds. Consumer Behavior and Marketing Management. Boston: Houghton Mifflin Company, 1967.
- Myers, John G. Consumer Image and Attitude. Berkeley: Institute of Business and Economic Research, Graduate School of Business Administration, University of California, 1968.
- Newman, Joseph W. (Ed.) On Knowing the Consumer. New York: John Wiley and Sons, 1966.
- Nicosia, Francesco M. Consumer Decision Processes: Marketing and Advertising Implications. Englewood Cliffs: Prentice-Hall, 1966.
- Nicosia, Francesco M. "Consumer Research: Problems and Perspectives," Journal of Consumer Affairs, 3 (Summer 1969), 9-25.
- Nowak, Kjell, Benny Carlman, and Karl-Erik Wärneryd. Masskommunikation och åsiktsförändringar. Stockholm: Kungl. Boktryckeriet P.A. Norstedt & Söner, 1966.
- Perloff, Robert. "Consumer Analysis," Annual Review of Psychology, 19 (1968), 437-66.
- Pollay, Richard W. "A Model for Research Planning in Consumer Behavior," in M. Venkatesan (Ed.) Proceedings. 3rd Annual Conference. Iowa City: Association for Consumer Research, 1972, 594-601.
- Robertson, Thomas S. Consumer Behavior. New York: Scott, Foresman and Company, 1970.
- Robertson, Thomas S. Innovative Behavior and Communication. New York: Holt, Rinehart and Winston, 1971.
- Robertson, Thomas S. and Scott Ward. "Consumer Behavior Research: Promise and Prospects," in Scott Ward and Thomas S. Robertson (Eds.), Consumer Behavior: Theoretical Sources. Englewood Cliffs: Prentice-Hall, 1973a, 3-42.
- Robertson, Thomas S. and Scott Ward. "Toward the Development of Consumer Behavior Theory," in B.W. Becker and H. Becker (Eds.), Combined Proceedings 1972. Chicago: American Marketing Association, 1973b, 57-64.
- Rogers, Everett M. and F. Floyd Shoemaker. Communication of Innovations. New York: The Free Press, 1971.
- Sandell, Rolf G. Att första konsumenten. Stockholm: P.A. Norstedt & Söners Forlag, 1969.

- Schlinger, Mary J. (Ed.) Advances in Consumer Research. Vol. 2. Chicago: Association for Consumer Research, 1975.
- Sheth, Jagdish N. "A Review of Buyer Behavior," Management Science, 13 (August 1967), 718-56.
- Sheth, Jagdish N. (Ed.) Models of Buyer Behavior: Conceptual, Quantitative, and Empirical. New York: Harper and Row, 1974.
- Sheth, Jagdish N. "The Next decade of Buyer Behavior Theory and Research," in Jagdish N. Sheth (Ed.), Models of Buyer Behavior: Conceptual, Quantitative and Empirical. New York: Harper and Row, 1974, 391-406.
- Sommers, Montrose S. "Problems and Opportunities in the Development of Consumer Behavior Theory," in Robert J. Holloway, Robert A. Mittelstaedt, and M. Venkatesan (Eds.), Consumer Behavior: Contemporary Research in Action. Boston: Houghton Mifflin Company, 1971, 14-22.
- Sommers, Montrose S. and Jerome B. Kernan (Eds.) Explorations in Consumer Behavior. Austin: Bureau of Business Research, University of Texas at Austin, 1968.
- Tigert, Douglas J. "A Review of V. Parker Lessig, Personal Characteristics and Consumer Behavior: A Multi-dimensional Approach," Journal of Marketing Research, 9 (May 1972), 234-6.
- Tucker, W.T. "Consumer Research: Status and Prospects," in Reed Moyer (Ed.), Changing Marketing Systems...Consumer, Corporate and Government Interfaces. Proceedings of the 1967 Winter Conference of the American Marketing Association. Chicago: American Marketing Association, 1968, 267-9.
- Twedt, Dik Warren. "Consumer Psychology," Annual Review of Psychology, 16 (1965), 265-94.
- Venkatesan, M. (Ed.) Proceedings. 3rd Annual Conference. Iowa City: Association for Consumer Research, 1972.
- Walters, C. Glenn. Consumer Behavior: Theory and Practice. Revised Edition. Homewood: Richard D. Irwin, 1974.
- Walters, C. Glenn and Gordon W. Paul. Consumer Behavior: An Integrated Framework. Homewood: Richard D. Irwin, 1970.
- Ward, Scott. "Consumer Socialization," Journal of Consumer Research, 1 (September 1974), 1-14.
- Ward, Scott, and Thomas S. Robertson (Eds.) Consumer Behavior: Theoretical Sources. Englewood Cliffs: Prentice-Hall, 1973.
- Ward, Scott and Peter Wright (Eds.) Advances in Consumer Research, Vol. 1. Urbana, Illinois: Association for Consumer Research, 1974.
- Zaltman, Gerald, Christian R.A. Pinson, and Reinhard Angelmar, Metatheory and Consumer Research. New York: Holt, Rinehart and Winston, 1973.

## BEHAVIOR CONTROL: ARE CONSUMERS BEYOND FREEDOM AND DIGNITY?

Rom J. Markin and Chem L. Narayana  
Washington State University

### Abstract

Humanists as well as consumer welfare advocates have raised with alacrity alarm about "mind control," "hidden persuaders" and subconscious manipulators of motivation. Now as a science of behavior as developed by Skinner, Bandura, and a host of disciplined, highly credentialed scholars emerge, and as their methods demonstrated in the laboratory continue to be replicated throughout the social and commercial world, the question and the issue become at once more substantive and real. This paper explores this new "science" of behavior modification and seeks to trace the significance of such methods on consumer welfare.

### Introduction

In 1972 B. F. Skinner, the noted Harvard behaviorist-psychologist, published his now famous book, Beyond Freedom and Dignity.<sup>1</sup> Skinner argued that free will, the basis of consumer sovereignty and freedom of choice, is no more than an illusion, a fiction of man's vanity, and that behavior, instead, is controlled by external influences. Although often haphazard and uncontrived, control is sometimes deliberately exercised by men and institutions for the purpose of getting others to do their bidding. Skinner is, of course, in favor of control and believes that the survival of our culture can only be assured if individuals are conditioned to want those things that serve the group interests.

Skinner contends that human behavior can be predicted and shaped through behavioral engineering precisely as if it were a chemical or biological process. This developing science of control aims to change the environment rather than the people. Behavioral engineering focuses on attempts to alter observable actions rather than feelings through a unique method of conditioning that has been used with rather startling success in both laboratory animals and humans in various controlled settings or environments. Such elaborate details have been undertaken to control or contrive the environmental setting that the technology, i.e., the hardware, the computers, and the reward and reinforcement contingencies, is referred to as behavioral technology or psychotechnology. Nonetheless, the lynch pin of the method consists simply of giving rewards to mold the subject to the experimenter's purpose.

Perry London, an articulate authority on behavioral engineering asserts that until recently attempts toward control focused on coercion, persuasion, inspiration, or education, and were almost always unpredictable. He goes on to add, however, that:

All this is changing now, and means are being found, in all the crafts and sciences of man, society and life, that will soon make possible precise control over much of people's individual activities, thoughts, emotions, moods and wills. Never in human history has this occurred

<sup>1</sup>B. F. Skinner, Beyond Freedom and Dignity (New York: Alfred A. Knopf, 1972).

before except as fantasy.<sup>2</sup>

Behavior modification and control have mushroomed into prominence during the last decade and are being used with growing confidence throughout the population--from mental institutions to the marketplace, in children as well as adults.<sup>3</sup> Three factors among others have led to this dramatic and, for some, shocking capability.

1. Knowledge of learning principles. Advances in how and why people learn have been rapid in recent years. Here also the behaviorist with his focus upon observable actions rather than mentalistic or mind states has spearheaded the development and understanding of how people learn and, hence, how behavior repertoires are changed via the now known rules of reinforcement.

2. Proliferation of technology. The widespread adoption and availability of modern technology, i.e., such extensions of man as radio, television, movies and electronic devices make it possible to present almost anything we want taught in effective packages which guarantee attention and interest. Thus, communication is assured in ways which are both intelligible and memorable.

3. The widespread use of computers. The computer enables the behavior modifier to process and store massive amounts of information. Large computer stored data banks provide almost limitless possibilities for possessing insight into the controllee's past behavior, his activities, his interests, and his opinions.

### Marketing and Consumer Behavior

#### Understanding or Control?

Consumer analysts and researchers generally are loath to acknowledge that they wish or have a capability to control consumers. Studying consumers however is to attain understanding; understanding usually leads to a capability to predict; and a capability to predict increases the capability to control. But this capability worries the marketing practitioner as well as the consumer behaviorist because control is anathema to our concepts of consumer sovereignty and freedom of choice.

<sup>2</sup>Perry London, Behavior Control (New York: Harper and Row Publishers, 1969), 4.

<sup>3</sup>There are a number of very readable pieces of literature which extend and elaborate on this argument. In addition to those already cited, i.e., the work of B. F. Skinner and Perry London, the reader might also enjoy examining the arguments found in the following: Albert D. Biderman and Hubert Zimmer, The Manipulation of Human Behavior (New York: John Wiley and Sons, Inc., 1961) and William Hunt, Human Behavior and Its Control (Cambridge, MA: Schenkman Publishing Company, Inc., 1971).

Marketers no doubt want the power to control and in some instances they have the power and use it; yet they deny its effectiveness by asserting that the consumer has the autonomy or the choice to accept or reject whatever is being promoted or merchandised.

#### What Are "Freedom" and "Dignity?"

The words freedom and dignity both are ranked high in the lexicon of consumer behavior. Marketers invariably invoke these concepts in discussions relating to consumer sovereignty, especially when questions arise relating to control of market situations. It is, furthermore, an endemic feature of our national political and social heritage that we extol the virtues of freedom and dignity by defining freedom in terms of states of mind, attitudes, or feelings. Man's struggle for freedom is not at all due to a will to be free but instead to certain behavioral processes characteristic of the human organism, the chief manifestation of which is the avoidance or escape from aversive or undesirable features of the environment.

Our tendency is to attribute freedom to an individual when the contingencies that guide his or her behavior are not readily apparent. Hence, the consumer in the marketplace may appear to be acting autonomously whereas in reality that consumer is behaving in accordance with well designed and deliberately arranged contingencies. Consumers feel that they have decision freedom--that is a real choice--when the options among which they choose are about equally attractive, that is when the options offer fairly equal net gains after calculating the costs and payoffs.

Traditionally we have extolled freedom because the need was felt to free people from aversive control. To free the individual from exploitation, from despots and charlatans who controlled through punitive methods, it was necessary to convince individuals that they could be free or that the power that was being used against them derived from them. Such a behavioral process has led to the conclusion that all control is wrong. Too often we don't recognize the fact that we are also controlled when we do what we want. "When I can choose from among options, I have freedom," is the way many would define freedom. Yet we are unprepared for the control that is exerted in making one want to do what one does.

Somewhat easier to deal with, dignity is usually related to self worth measured in vague units of responsibility. We consider people responsible for what they do. If the causes for a person's behavior are conspicuous we credit the causes, but if the causes are inconspicuous we credit the person. In the case of consumer behavior we have characteristically understood so little about the causes, motives, reasons, or contingencies that produce that behavior that we credit the causes to the consumer. Thus the literature of dignity is concerned with preserving due credit. Marketers and consumer psychologists go on imputing both freedom and dignity to consumer buyers by giving them credit for what they do--the items they choose, the stores they select, their allegiance to certain brands. By so doing we continue to pay homage to the autonomous man when, more realistically perhaps, we should be examining the environment that is responsible for this putative worthy behavior.

#### Autonomous Man Versus Machine Man

Characteristically, the conventional wisdom asserts that man initiates, originates, and creates, and to this

extent his behavior is somewhat divine or autonomous.<sup>4</sup> Such a notion contends that the inner man wills an action but that the outer man executes it. Skinnerians argue that the concept of autonomous man is an anachronism. A science of behavior replaces the notion of autonomous man--the fictional idea that behavior is the result of inner feelings, states of mind, expectancies and all that--with the more modern notion that man is an elaborate machine. The behaviorists assert that we cannot account for the behavior of any system while staying wholly inside it. Eventually we must look at those forces acting upon the organism from without.

A machine model of man has sterile and repulsive aspects for many persons. Our traditional philosophy and Christian backgrounds refute the notion of man as machine. The machine model view does not mean, however, that man rusts, clanks, whirs, or goes chug-chug. Nor is he a large warm, soft computer. What it does mean in a scientific sense is that man's behavior--like the machine--is lawful and limited. Lawful behavior is predictable and the general principles of human behavior apply to all individuals.

Most scientists today view man as a kind of machine--hence John Reiner describes man as an "Environmentally Modifiable Physico-Chemical Regulatory Device."<sup>5</sup> George Miller alludes to man as "an information-processing and information gathering device."<sup>6</sup> Norbert Weiner defines a machine about like Miller does a man, as "a device for converting incoming messages into outgoing messages."<sup>7</sup> Finally John Von Neuman circumscribes man with the phrase "self reproducing automata."<sup>8</sup>

It must certainly be recognized that in relation to behavior, it is the doctrine of free will or autonomous man that is most threatened by the machine model of man. The fact is, most Skinnerians contend, actions and behavior are determined by the environment--not by the inner man, mind states, mentalistic feelings, attitudes, perceptions, or other such fictions.

#### What Causes Behavior?

Current use of the terms cause and effect involve relating a cause to an independent variable and effect to the dependent variable. In consumer behavior models following this approach, the independent variable has been attributed to such mechanisms as attitude, perceptions, motivations, or other internal mind states; or internal drives such as achievement, affiliation, and aggression. Because these internal traits cannot be observed, their presence must be inferred from the actual observation of objective behavior. Thus if a person

<sup>4</sup> B. F. Skinner, Science and Human Behavior (New York: The Free Press, Collier Macmillan Limited, 1953). This book is a somewhat more technical treatment than Beyond Freedom and Dignity. However, it too is written at the intelligent layman's level of comprehension.

<sup>5</sup> John M. Reiner, The Organism as an Adaptive Control System (Englewood Cliffs, NJ, Prentice Hall, Inc., 1968).

<sup>6</sup> George Miller, The Psychology of Communication (New York: Basic Books, Inc., 1957).

<sup>7</sup> Norbert Weiner, God and Golem, Inc., A Comment on Certain Points Where Cybernetics Impinge Upon Religion (Cambridge, MA: MIT Press, 1964).

<sup>8</sup> John Von Neuman, Theory of Self Reproducing Automata, Arthur Bucks (ed) (Urbana, IL: University of Illinois Press, 1964).

demonstrates a high degree of brand loyalty to a particular brand, we infer that he holds a favorable attitude toward the brand. We then generalize that he is brand loyal to Product X because he has a favorable attitude toward Product X. Notice that the reasoning is circular.

A prevailing modern point of view attributes the causes of behavior to goals or purposes. Thus instead of arguing that behavior is caused by drives which push, we assert that behavior is caused by goals which pull. Too often the practice of looking inside the individual for an explanation of behavior has tended to obscure the variables outside the organism which are readily observable for direct scientific analysis. The fact is, state the behaviorists, behavior is not determined from within but from without--behavior is shaped by its consequences. By setting up contingencies of reinforcement, a particular bit of desired behavior is rewarded to make sure that it will be repeated.

#### Modifying Behavior--Insight or Action Approaches

Our young discipline of consumer behavior has characteristically assumed that behavior is attributable to inner states of mind. This insight approach to consumer behavior would posit that motives, attitudes, cognitions, or other psychological causes dictate behavior. The marketer who uses the insight approach attempts with his advertising to create favorable attitudes, cognitions, or images toward his product or institution; and these favorable inner psychological mind states are then supposed to increase the consumer's predisposition for the seller's products. It is the old Aristotelian idea that "what is impressed will be expressed." Marketers who use insight techniques are always suggesting that their products will make the consumer happier, more socially acceptable, more poised, or generally more successful. Insight approaches used in both print and electronic media are most often straightforward projections of information or ideas designed to lead customers to an understanding of themselves and their need systems through the use of association, interpretation, and perceived relationships. The copy in print media insight advertising is usually a statement of claims and purposes. For example, an Excedrin advertisement states:

"You might guess just by looking that Excedrin has more pain reliever than the aspirin tablet. But more importantly, there is evidence from two medical research studies that Excedrin performs significantly better than the common aspirin tablet. More pain reliever. And evidence of its greater effectiveness. That's what Excedrin has to offer you."

In television advertising which utilizes insight approaches, the presentations are equally straightforward and the message is projected as a monologue by the announcer.

More bizarre and less frequently employed insight approaches are often based upon the notion that the inner causes or motives which compel people to buy goods are centered deep within the emotions or personality of the individual. These hidden emotions are often thought to border on libidinous themes and unusual symbolism. Ferreting out these emotions is the task of the motivation researchers who use small samples and subjective techniques borrowed from the psychoanalytic school.<sup>9</sup>

<sup>9</sup>"Psyching Them Out: Ernest Dichter Thrives Selling Firms Research on 'Hidden Emotions.'" Wall Street Journal, (Monday, November 20, 1972), 1.

However, determining what selling and merchandising appeals based upon these findings to direct to consumers has never been a question easy to resolve.

It would appear that more and more advertisers are increasingly moving to action oriented approaches to behavior modification. In some instances their use is not so much by design as the result of happy or fortuitous accident. Instead of focusing on the motives, attitudes, or mentalistic states that allegedly produce a consumer's behavior, i.e., dissonance, search, evaluation, or information processing, action approaches tend to focus effort on the behavior proper without much concern over its origins or its meaning.

According to action oriented theorists, most consumer symptoms (anxiety, dissonance, perceived risk, search, confusion, overconsumption, and so forth) are really nothing more than habit patterns which have been learned or acquired through a process of conditioning and are therefore capable of being extinguished or modified through several techniques of demonstrated effectiveness. Action oriented advertisers assume their respective purposes in terms of specific behavioral objectives to be (1) the disclosure of problems, (2) the channeling of attention, and (3) the reformation of behavior.

Action approaches are comprehensive attacks beginning with massive data accumulation which provide baseline information on existing behavior. For this reason alone, action approach advertisers utilize heavily quantitative studies based upon large samples and extensive use of statistical analysis concerned with mapping consumer space preferences and psychographic and life style analyses. These studies provide telling data regarding the activities, interests, and opinions of market segments such as heavy users so that appeals which are compatible with their life styles can be developed and, more importantly, so that these appeals can be projected in a total managed setting to assure the greatest amount of exposure and interest in the given advertisement.

Behavioristic action oriented approaches are most concerned of course with the reformation of behavior and with the continued reinforcement of already generated desired behavior. Behavioristic advertisements attempt to simulate a problem oriented situation by dramatically illustrating the problem. For example, a bad cold, an upset stomach, a worker with aching muscles or one who needs something for a headache, a busy harrassed homemaker who needs help in the kitchen, children with too many activities, a husband who can manage the little expenses but doesn't want to face the prospect of complete financial disaster in the case of major sickness or other health emergencies, an older man who is cross or out of sorts with his grandchildren because of irregularity. Once the problem is dramatized as real and relevant, the solution is posed. The advertisement invariably stresses the way practical and sensible people--"people like you and me"--solve the problem. Finally, the rewarding, reinforcing, satisfying aspects of the solution are strongly emphasized.

The emphasis on action oriented approaches can be greatly attributed to the increasing use of television as an advertising medium. Television affords a unique opportunity to arrange desired contingencies. It provides an astounding number of perceptual sensations; color, brightness, size, shape, movement, volume, and pitch. Plus, television has so great a potential as an attention getting medium, an affect-creating medium and, of course, as a message-creating medium. Hence, the most successful of action oriented advertisements are frequently television commercials. Proctor and Gamble has reaped unusual successes utilizing this approach with such products as Head and Shoulders shampoo, Ivory Liquid detergent, and most remarkably with Crest tooth-

paste. Such commercials usually involve a transaction or dialogue, between two or more persons. One of these persons invariably is shown complimenting or stroking the problem solver for his or her adeptness, concern, practicality, or sagacity. This device is easily recognizable in such well known ads as Crest's "Look Mom, no cavities;" Nyquil and Absorbine Jr., "You're a good wife Mabel;" Geritol's "How do you stay so young and beautiful Mother?" Jello's "You're a good Mom;" Pillsbury's "Nothing says loving, like something from the oven." This technique is nothing other than operant conditioning which consists of reinforcing or rewarding consumers, in a simulated fashion, for appropriate responses. Many of today's most successful products are promoted and advertised on the basis of such action approaches--Coca Cola, Pepsi Cola, McDonalds Drive-in, Kentucky Fried Chicken, Nyquil, Absorbine Jr., Alka Seltzer, Phillips Milk of Magnesia, Pepto Bismol, Folgers coffee, Crest toothpaste, Head and Shoulders shampoo and many other products. In some instances the appeal and contingency management are overt or even blatant. In others the appeal and the staging are subtle and indirect. Action approaches work because of vicarious modeling. People have a tendency to emulate those things that bring satisfaction to others in the hope of securing similar rewards and satisfaction for themselves.

Advertisers cannot animate consumer puppets through operant conditioning, but they can shape behavior and increase the probability of generating desired responses. And marketers and behavior modifiers in the world of commerce are beginning to look increasingly at these action approaches as a means of directly controlling consumer behavior--either as single strategies or in conjunction with the older and more prevalent insight approaches.

#### The Pertinent Question

A Pertinent question has been raised which suggests that we had best begin to debate seriously the merits of operant conditioning and action approach strategies in the marketplace. Hansen asks, "Are stable traits residing within the individual the major source of variation in behavior, or is such variation mainly ascribable to differences in the situations people encounter?"<sup>10</sup> Recent research tends to question the notion that behavior is a function of inner causes such as attitudes, personality, or motivations.<sup>11</sup> And furthermore the reason which is offered in explanation of this condition is that such factors as personality or attitudes fail to take into account important situational variables. Our young discipline of consumer behavior has probably been too preoccupied with studying consumer traits, internal mind states, and subjective unobservable mechanisms such as personality, attitude, motives, and perceptions, and too little concerned with how situations or environments might be manipulated in order to modify behavior. Our models of consumer be-

<sup>10</sup> Fleming Hansen, Consumer Choice Behavior: A Cognitive Theory (New York: The Free Press, 1972), 42.

<sup>11</sup> In the case of criticism concerning personality studies the reader might like to examine the article by J. M. Hunt, "Traditional Personality Theory in the Light of Recent Evidence," American Scientist, 53 (1965), 80-96. Attitude studies are somewhat more numerous. For example see, R. A. Bauer, "Does Attitude Change Take Place Before or After Behavior Change?" in L. Adler and I. Crespi (eds) Attitude Research on the Rocks (Chicago, American Marketing Association, 1968). Also, H. J. Ehrlich, "Attitudes, Behavior, and the Intervening Variables," The American Sociologist, Vol. 4, No. 1, (February 1969), 29-34.

havior are almost all cognitive models which stress the reasoning, information processing, mentalistic aspects of behavior. While some little attention is devoted to the issue of how the consumer transforms the situational input to behavioral output, the consumer is still too often viewed in terms of inner drives and mind states; as the autonomous man who originates and creates rather than as the organism who reacts, responds, and is shaped and modified by the astute manipulation of his environment or by clever situation management. Management is essentially planning, and planning is concerned with controlling the events in one's environment--that is the essence of contingency management.

#### Toward A "Science" of Behavior

The behaviorists contend that a "science of behavior" must replace the intuitive wisdom of the old style diagnostician with the objective and analytical procedure of the clinic. And when such procedures lead us to an understanding of the laws and generalities of a system, we are then ready to deal effectively with some part of our world. By predicting the occurrence of an event we are able to anticipate and cope with that event. Furthermore, by arranging and controlling conditions in the ways suggested by the laws of a system we then not only anticipate and predict, we control or cause an event to occur or to assume certain dimensions. Such is the behaviorist notion of a science of behavior.

Behaviorism is primarily dichotomized into two schools; the classical conditioning school and the operant conditioning school. Both classical conditioning via association and operant conditioning are important doublebarreled approaches to scientific behavioral engineering. Conditioning is the basis of nearly all habits, all skills and in some people's estimation, nearly all learning. Classical conditioning is the cornerstone of behavioral engineering because through classical conditioning and association we can control and modify the emotions, especially mood. And mood is the emotional basis of behavior; manipulating one tends to coerce the other.

Operant conditioning means the learning of instrumental behavior, i.e., behavior that serves a purpose, solves a problem, answers a question, provides escape from aversion or pain, or contrariwise leads to the attainment of pleasure. Most behavior, at least a large part of consumer behavior, is trial and error involving several responses. The most useful responses are habituated or learned. When the consumer acts or behaves in the marketplace, his behavior is goal or problem oriented. If a product satisfies him, he is reinforced, and the probability that he will purchase the product subsequently is increased. If on the other hand, his behavior is punished or negatively reinforced, he continues to seek new responses by engaging in alternation behavior until such time as his responses are more suitably rewarded. Thus the consumer is an operant, and through operant conditioning we can shape and modify almost all instrumental or purposive actions. If behavior is shaped by its consequences we need only to control and to manipulate the consequences in order to manipulate or shape the behavior.

Hence through the dual application of classical conditioning and operant conditioning we may well be moving toward a science of behavior that is not only deliberate but equally precise. Such a science must manifest four capacities:<sup>12</sup> (1) to produce a specific variety of effects, (2) to control the intensity of effects, (3) to

<sup>12</sup> London, op. cit., 25ff.



specify the domain of effects, and (4) to control the duration of effects. Right now these conditions can be met in the laboratory and the raw material for behavior control in the laboratory has become the catalytic agent in the marketplace. The basis for a behavior technology in the marketplace is simply a plan for implementing what is already known. Although these raw materials and capabilities have already been used casually, perhaps even by accident, their planned systematic use is increasing and eventually they will be used more and more deliberately.

#### How Behavior is Controlled and Modified

In spite of the really great differences and wide variations exhibited in consumer behavior, there are at the same time great similarities in behavior manifested by those whose past experiences and current situations are similar. Hence their behavior is obviously shaped by their similar environments and these environments are influencing because of such phenomena as socialization, modeling, and culture.

Socialization. Socialization and the socialization process have to do with the activities and methods by which people influence one another through mutual interchange of thoughts, feelings, and actions. In so many words we are taught to value and esteem particular actions. Such reinforcement leads to the repetition of these actions; thus through the process of social learning we are taught the ways of our society in relation to our own need fulfillment activities. Hence socialization is concerned with behavior modification and control and it utilizes the technology and procedures of operant conditioning such as social reinforcement, praise, instruction, and prompting.

Modeling. Modeling is a particular and unique phenomenon related to socialization. Evidence supports the modeling theory of behavior modification to the extent that what one sees guides and influences what he does. The behavior of both children and adults is shaped and modified by such psychological processes as social perception, interpersonal perception, and the social exchange process of interaction. These phenomena, basically concerned with social learning, ultimately shape and modify the response repertoire. For example, television has become potentially, if not realistically, a more important means for facilitating modeling than any other medium or institution. A steady diet of television has a powerful influence on the viewer.<sup>13</sup> Television shapes the way people view the world and the kind of people they will be. We can let it shape others as it sees fit or we can shape it to fit our notions of what the culture ought to be.

Opinion leaders, key influentials and significant others are all instrumental in shaping consumer behavior. Such is the essence of social exchange theory. If receiving approval or rewards is reinforcing to someone, his behavior can conceivably be shaped by withholding the approval or rewards and making them contingent on appropriate or desired behavior.<sup>14</sup>

Culture. Culture is, in effect, learned behavioral

<sup>13</sup>For a more detailed view see Scott Ward and Daniel B. Wackman "Family and Media Influences on Adolescent Consumer Behavior," American Behavioral Scientist, 14(Jan-Feb., 1971), 415-427.

<sup>14</sup>The notion of social exchange theory is developed more extensively by George C. Homans, Social Behavior: Its Elementary Forms (New York: Harcourt Bruce Jovanovich, 1961).

dispositions passed from one generation to another--a form of nonbiological heritage. The social environment, or culture, shapes and maintains the behavior of those who live in it by establishing the rules, norms, and expectancies of behavior. Hence our children are taught to mind Mommy and Daddy, to love their country, to respect God, to brush their teeth (preferably with Crest), to eat their vegetables, and to do whatever else the cultural norms, customs, and folkways dictate at a particular time.

Certainly our culture, consumption and materialistically oriented, affects and shapes consumer behavior. We are still largely work ethic oriented because work is the means to money and money is the means to goods. Material well being--the ability to buy washing machines and corn poppers--is much of what America has always been about. Americans are not born, however, with a drive for goods--this capacity is shaped and influenced by the very nature of the materialistic-acquisitive orientation of the culture.

#### Design or Accident

As long as behavior shaping and control via such phenomena as socialization, modeling, and culture are perceived as loose jointed, uncoordinated efforts of benevolent controllers, we view the impact of such control as accidental, innocuous, and therefore nonthreatening. However, if these efforts are seen in the light of a grand design by a malevolent or totalitarian controller they are likely to be construed as vicious, dangerous, subversive, and evil. Our perceptions of these two opposing viewpoints are shaped by our own past experiences through socialization and inculturation. We evaluate alien attempts to control as coercive, whereas our own efforts, perhaps because they are more subtle, are not considered coercive. Our western notions condition us to believe that control via information is hardly control at all. Yet education, prayer, rhetoric, propaganda, demagoguery, semantic seduction, and advertising are all typical examples of control by information.

The elicitation of personal information about someone increases enormously the possibility for controlling him. And we are increasing with geometric proportions our knowledge about people--their attitudes, their behavior, and their innermost feelings about the most private aspects of their lives such as sex, finances, religion, and so on. The use of data banks, psychographic profiles, and perceptual mapping devices have effectively extended our range for controlling behavior.<sup>15</sup> This increased knowledge about consumer's, voter's, and viewer's activities, interests and opinions can be used effectively by a potential controller to manage the contingencies so that the environment of behavior is more pleasant and reinforcing and thus more likely to generate the desired kind and frequency of behavior. The expenditures for voter analysis, media studies, and market research are in the interest and design of a better system of control. Information is knowledge and knowledge is power -- over people.

<sup>15</sup>The significance of each of these respective procedures is detailed in the following:

Richard L. Nolan, "Computer Data Bases: The Future is Now," Harvard Business Review, 51-5(Sept-Oct., 1973), 98-115.

William D. Wells and Douglas J. Tigert, "Activities, Interests and Opinions," Journal of Advertising Research, 4(Aug. 1971), 27-35.

P. E. Green and F. J. Cormone, "Multidimensional Scaling: An Introduction and Comparison of Nonmetric Unfolding Techniques," Journal of Marketing Research, 6-3 (Aug., 1969) 330-341.

Consumer behavior is largely affected and shaped by operant conditioning activities. What is called seduction by intellectuals is called salesmanship and marketing by most other people. The chief means of consumer behavior control and modification is information dissemination. Most of this information is verbal and the prototype for verbal control is again salesmanship and marketing. However, marketers affect and shape behavior by means other than verbal cues. Product development, styling, and design; promotional activities such as advertising, personal selling, testimonials, and modeling; pricing, store location, interior store design, layout, spatial arrangement, decor, color, and lighting; all of these modern merchandising devices--methods for structuring the environment of buying and consumption--are control efforts for shaping behavior through rewards and reinforcement.

Is it just a fortuitous accident or is it primitive psycho-logic that merchandise is so frequently called "goods?" Perhaps it is both, but "good" is a powerful reinforcer and goods as merchandise and services are increasingly sought for their rewarding and reinforcing effects. If behavior is shaped by its consequences, consumers are wooed and rewarded with a thousand gew-gaws, gadgets, praises, and approbations. They are told to "indulge themselves," that they "deserve a break today," that "they only go around once," to enjoy "the real thing"--in short in an increasingly hedonistic culture to seek pleasure--and the marketplace is the place where pleasure is found.

This is not to suggest that consumers are mindless automatons. Human behavior is a function of millions of nerve cells with a multi-complex spatio-temporal integration of so many factors that its direct control and manipulation is not feasible. In particular, high-level, top of the head, rational-cognitive decision processes are not directly manipulable.

We should not hurriedly assume, however, that really important or seemingly sophisticated consumer behavior is therefore safe from the effects of shaping and control. If anything, what we need badly to acknowledge is that a large portion of consumer behavior is not cognitive or rational to begin with but rests, not in the higher reaches of the intellect, but in the middle, back, and sides of the head, and in the anxieties, hostilities, lusts, and frustrations which humans in general and consumers in particular share with lower animals.

The fact that so much consumer behavior is shaped in an operant fashion perhaps even refutes the notion of the cognitive consumer; the model of a highly rational goal striving, problem solving, information seeking, extracting, processing consumer; a consumer who consciously reasons; a prowling, roving human computer. Such a traditional concept may very well reflect a tragic misunderstanding of our notions regarding consumers' intelligence, rationality and intellectual autonomy. Empirical research on consumer decision processes has, in fact, documented that consumers:

1. Do not seek extensive amounts of information in relation to purchase and consumption problems
2. Do not process large amounts of information in relation to purchase and consumption problems
3. Do not appear to engage in extensive problem solving behavior even in relation to big ticket or capital intensive items such as automobiles, houses, and

major appliances.<sup>16</sup>

It is also rather well known but only begrudgingly acknowledged in the formal literature of consumer behavior that most decisions are made on the basis of limited cognitive activity involving selective cues and that these cues are more in the psychological realms of the effective-emotional amygdala than in the cognitive realm of the cerebral cortex.

Thus a more relevant model of the consumer would be one possessed of more realistic attributes; attributes which acknowledge the frailty of the human condition. This model would admit to the emotional-affective nature of the consumer. Consumer behavioralists, much like Hamlet, are too prone to look at man and say, "How like a god!" Perhaps they should, like Pavlov, take another look and say, "How like a dog!"<sup>17</sup>

Behavioral engineering is practiced upon consumers because those who attempt to influence consumers have power. The giant corporations of America constitute the planning system for engineering and programming public tastes.<sup>18</sup> Rather than being governed by market forces, these corporations control the market by setting prices and creating consumer demand for products. Those who manage the great corporations that comprise the planning sector of our economy have vast manipulative and shaping power. They develop their own peculiar purposes and they are able to impose their own purposes on others which is the very basis of power. The technost-structure of this corporate sector consists of multi-level merchandisers such as the big 8 oil companies. Secrecy is a passion within these companies because corporate managers understand better than anyone else that controlling information is the key to exercising power. Of course, these demand and taste shaping corporations play an interesting politico-psychological game. They recognize that the safest way to exercise power is to pretend that it resides elsewhere. Thus, the corporate multi-level merchandiser likes to dwell on the notion of the sovereign customer who dictates his will to the obedient producers of electric forks, multicolored underarm deodorants and folding water beds. This is, of course, consistent with their complementary pronouncements which assert that the corporation is socially responsible, subordinate to the state, and that the anti-trust laws are adequate to preserve the market.

Thus all "people fixers" share a common purpose--to mobilize, repair, and program man. Behavioral engineering is alive and well in consumer behavior because we have a consumer who is receptive and merchandisers and corporate planners with the power and the inclination.

#### The Ethics of Behavior Control: A Summary and Conclusion

Behavior control techniques are neither good nor bad as such, but it would go without saying, they can be misap-

<sup>16</sup>These ideas and the research upon which they are predicted can be reviewed more extensively in Rom J. Markin, Consumer Behavior: A Cognitive Approach (New York: The Macmillan Company, 1974), especially Chapter 17.

<sup>17</sup>The analogy is only metaphorical and suggests that human organisms are largely emotional-affectively oriented. Even a dog can say "I like it." This implies no demeaning of the human condition.

<sup>18</sup>These ideas are based upon those of John Kenneth Galbraith, Economics and the Public Purpose (Boston: Houghton Mifflin Company, 1973).

plied. It would seem that we need behavioral accountants as well as behavioral engineers. That is, we need responsible people to weigh, analyze, and measure the impact of behavior shaping and to evaluate its effects in terms of the overall welfare of the culture. Even this suggestion has frightening portents for some. There is such a well ingrained aversion to the ideology of control that we are all afraid to manipulate people. We live in a twilight of authority. We accept the status quo without qualm or reaction. We are so afraid of effective controls that we go right on accepting controls that are subtle, insidious, and sometimes even malevolent and injurious.

Our western humanistic philosophy has taught us that control presumes to relegate man to an animal status, robbing him of his choices and foreclosing on his possibilities and potentialities. We believe in a noncoercive control system where people are not seduced into compliance, yet our entire marketing-consumption oriented economy is based upon seduction or sales promotion. Our traditions have further shaped us to believe in self direction--the idea that people should be free, to decide for themselves how they want to live their own lives. Yet again our belief systems are prostituted when compared with our practices. We mouth these notions, yet as educators, administrators, marketers, and governmental officials, we continue to talk about planning, organizing, controlling, contingency management, and situation management.

Consumer behavioralists are not immune from this hypocrisy either. Most of us have had a kind of schizoid attitude regarding what it is we are about, our intentions, our motives, and just who it is we are attempting to serve. Some small part of the time we act as if we are attempting to create a better environment for fish but our real efforts are more in the nature of building and designing better hooks for catching more fish. Thus, we too pay homage to the myth of freedom and dignity but we are not opposed to subjecting consumers to the indignity of the Skinner box.

Consumer behavioralists need to teach people how to be better consumers, not necessarily how to consume more. We need to teach people how to use and seek information; how to sharpen their perceptual learning abilities. The alternative to control is awareness. An aware, enlightened, knowledgeable consumer is better insulated against manipulation. The consumer movement itself is an attempt to free the individual from aversive control and to assure the consumer's right to safety, to be informed, to choose, to be heard.

The most effective means of controlling the controllers is through counter control. For without counter control, the controller has an increasing tendency to become first careless, then callous, and ultimately cruel. Through counter control efforts, an enlightened consumer or an honest, creditable, judicious government can bring about a balance of power between the market planners and consumers. Power does corrupt but the only antidote to power is power and the best measures for counter control rest with the power of government, education, families, business, and individuals. And this power is the power to know, to reason, to seek, to ask, and to demand response and initiative.

Control seen in this perspective is not quite so awesome in its prospects. Even though man is controlled largely by his environment, he, in turn, largely controls his environment. It is only through awareness and through deliberate efforts at counter control, plus courage and insight, that we may be able to recapture from behavioral engineering and technology both our freedom and our dignity.

## References

- R. A. Bauer, "Does Attitude Change Take Place Before or After Behavior Change?" in L. Adler and I. Crespi (eds) Attitude Research on the Rocks, (Chicago: American Marketing Association, 1968).
- Albert D. Biderman and Hubert Zimmer, The Manipulation of Human Behavior (New York: John Wiley and Sons, Inc., 1961).
- H. J. Ehrlich, "Attitudes, Behavior, and the Intervening Variables," The American Sociologist, 4 (February, 1969), 29-34.
- John Kenneth Galbraith, Economics and the Public Purpose (Boston: Houghton Mifflin Company, 1973).
- P. E. Green and F. J. Cormone, "Multidimensional Scaling: An Introduction and Comparison of Nonmetric Unfolding Techniques," Journal of Marketing Research, 6 (August, 1969), 330-341.
- Fleming Hansen, Consumer Choice Behavior: A Cognitive Theory (New York: The Free Press, 1972).
- George C. Homans, Social Behavior: Its Elementary Forms (New York: Harcourt Brace Jovanovich, 1961).
- J. M. Hunt, "Traditional Personality Theory in the Light of Recent Evidence," American Scientist, 53 (1965), 80-96.
- Perry London, Behavior Control (New York: Harper and Row Publishers, 1969).
- Rom J. Markin, Consumer Behavior: A Cognitive Approach (New York: The Macmillan Company, 1974).
- George Miller, The Psychology of Communication (New York: Basic Books, Inc., 1967).
- Richard L. Nolan, "Computer Data Bases: The Future is Now," Harvard Business Review, 51 (September-October, 1973), 98-115.
- "Psyching Them Out: Ernest Dichter Thrives Selling Firms Research on 'Hidden Emotions.'" Wall Street Journal, (Monday, November 20, 1972), 1.
- John M. Reiner, The Organism as an Adaptive Control System (Englewood Cliffs: Prentice Hall, Inc., 1968).
- B. F. Skinner, Beyond Freedom and Dignity (New York: Alfred A. Knopf, 1972).
- B. F. Skinner, Science and Human Behavior (New York: The Free Press, Collier Macmillan, Ltd., 1953).
- John Von Neuman, Theory of Self Reproducing Automata, Arthur Bucks (ed) (Urbana: University of Illinois Press, 1964).
- Scott Ward and Daniel B. Wackman "Family and Media Influences on Adolescent Consumer Behavior," American Behavioral Scientist, 14 (January-February, 1971), 415-427.
- Norbert Weiner, God and Golem, Inc., A Comment on Certain Points Where Cybernetics Impinge Upon Religion, (Cambridge: MIT Press 1964).
- William D. Wells and Douglas J. Tigert, "Activities, Interests and Opinions," Journal of Advertising Research, 4 (August, 1971), 27-35.

CORRELATES OF CONSUMPTION OF FOOD NUTRIENTS

BY U.S. URBAN HOUSEHOLDS

Mohamed Abdel-Ghany, University of North Carolina-Greensboro  
Gordon E. Bivens, University of Missouri-Columbia

Abstract

Nutrient and calorie intakes of U.S. urban households were examined in relation to six economic and socio-demographic characteristics. Adjusted Household Meal Units was most frequently, and positively, associated with nutrient intakes. Income was positively associated with five individual nutrients. Education and age of female homemakers tended to be negatively associated with nutrient and calorie intakes.

In recent years much attention has focused on the deterioration of the nutritional level of American diets. Despite the availability of sizeable and diverse food supply, some Americans consume less than recommended dietary allowances of food nutrients. Unequal distribution of economic resources in America has led to the existence of pockets of poverty and associated malnutrition; however, some high income households also suffer from malnutrition either because of lack of nutritional information or because of affluent life styles. This paper examines how households of differing socio-economic characteristics vary in their consumption of food nutrients.

Method

Sample

The source of data for this study was the U.S.D.A. Household Food Consumption Survey of 1965-66, the most recent nationwide survey of household food consumption. The sample was limited to the data collected in the spring of 1965 (the largest single wave of four seasons of interviewing); also the sample was confined to the urban black and white households. Households with missing responses for any of the variables used in the analysis were eliminated. Accordingly, 3,860 households constituted the studied sample.

The Dependent Variables

Designated as dependent variables were per household weekly actual consumption of food energy and nine nutrients -- protein, fat, calcium, iron, vitamin A, thiamin, riboflavin, niacin, and ascorbic acid.

The Independent Variables

The independent variables included: the education of female homemaker, household income after taxes, age of female homemaker, race of female homemaker, region of the country (Northeast, North Central, West, South), and Adjusted Household Meal Unit. The measure of Adjusted Household Meal Unit (AHMU) was devised to take into account differences among households in number of meals eaten at home by the various members as well as the household size and composition. The procedure for the derivation of the AHMU variable was as follows:

(1) consumption equivalence scales for 20 sex-age categories were computed based on the USDA moderate cost food plan (U.S. Department of Agriculture, 1965). The 20-35 year old male category is used as the base; the cost of one week's food for each of the other sex-age categories is divided by the estimated cost of one

week's food for a 20-35 year old male. The consumption equivalence coefficients for the 20 sex-age categories are presented in Table 1. (2) The total number of meals served each member of the household per week is multiplied by the consumption equivalence coefficient applicable for each household member's sex-age category. (3) The sum of the adjusted meals for all members of the household is calculated. The AHMU, then, is

$$\sum_{i=1}^n (\text{the total meals served per week to each household member} \times \text{consumption equivalence coefficient applicable for each household member})$$

where  $i$  = the number of household members.

Method of Analysis

Step-wise least squares multiple regression was used as

TABLE 1

Cost of One Week's Food Under the U.S.D.A. Moderate-Cost Food Plan and Consumption Equivalence Scale Coefficients for Different Sex-Age Categories, 1965

Sex-Age Category	Cost of One Week's Food	Scale Coefficient
<b>CHILDREN</b>		
Under 1 year	\$ 3.90	0.40
1-3 years	4.90	0.50
3-6 years	5.90	0.61
6-9 years	7.10	0.73
<b>GIRLS</b>		
9-12 years	\$ 8.20	0.84
12-15 years	9.00	0.93
15-20 years	9.20	0.95
<b>BOYS</b>		
9-12 years	\$ 8.30	0.86
12-15 years	9.80	1.01
15-20 years	11.20	1.15
<b>WOMEN</b>		
20-35 years	\$ 8.50	0.88
35-55 years	8.20	0.84
55-75 years	7.10	0.73
75 and older	6.30	0.65
pregnant	9.90	1.02
nursing	11.30	1.16
<b>MEN</b>		
20-35 years	\$ 9.70	1.00
35-55 years	9.00	0.93
55-75 years	8.30	0.86
75 and older	8.00	0.82

the technique of analysis since it permits examination of the effects of each independent variable on a given dependent variable while controlling for the simultaneous effect of other independent variables. This technique is capable of handling variables measured in either quantitative or qualitative terms. Within the regression models, qualitative variables appear as sets of dummy variables; these dummy variables are given a value of one if an observation falls in a designated category, zero if it does not.

The Statistical Model

The multiple regression model used to test the significance of the different household characteristics may be represented as follows:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_{41} + \dots + b_6X_{43} + b_7X_5 + b_8X_{61} + \dots + b_{10}X_{63}$$

where:

- Y is the dependent variable,
- a is a constant term,

- X<sub>1</sub> is the household income after taxes,
- X<sub>2</sub> is the adjusted household meal unit,
- X<sub>3</sub> is the education of female homemaker,
- X<sub>41</sub> .....X<sub>43</sub> is an expression for age of the female homemaker in the form of dummy variables,
- X<sub>5</sub> represent the race of the female homemaker in the form of a dummy variable,
- X<sub>61</sub>.....X<sub>63</sub> is the region variable in the form of dummy variables,

and

b<sub>1</sub>.....b<sub>10</sub> are the regression coefficients.

It should be noted that in testing for the significance of a dummy variable coefficient, we are testing to see if that coefficient is significantly different from that of the omitted category. For continuous variables, however, the coefficients indicate the direction and magnitude of the effect on the dependent variable of a one unit increase in the continuous independent variable.

TABLE 2

Summary of Associations Between Household Characteristics and Nutrient Consumption.  
 ('+' indicates positive association statistically significant at .05;  
 '-' indicates negative association statistically significant at .05)

Independent variable	Dependent variable									
	Calories	Protein	Fat	Calcium	Iron	Vitamin A	Thiamin	Riboflavin	Niacin	Ascorbic Acid
Adjusted Household Meal Unit	+	+	+	+	+	+	+	+	+	+
Household Income		+	+			+			+	+
Education of Female Homemaker	-	-	-		-		-	-	-	+
Age of Female Homemaker (25-44 year age group used as the comparison group)										
14-24 years									-	-
45-64 years	-	-	-	-	-		-	-		+
65 years or over	-	-	-	-	-		-	-	-	
Race of Female Homemaker (White households used as the comparison group)										
Blacks					-	+		-	-	
Region of the United States (South used as the comparison group)										
Northeast	-		-			+		+		+
North Central									+	
West			-			+				

## RESULTS

A summary of results is presented in Table 2 and details for each nutrient are reported in Tables 3 through 12. Overall, the percentage of variation in the household consumption of individual nutrients which was explained by the six independent variables was: energy, 61; protein, 59; fat, 55; calcium, 55; iron, 55; vitamin A, 16; thiamin, 58; riboflavin, 54; niacin, 53; ascorbic acid, 31. In the further discussion, it should be noted that the significance of the regression coefficients is measured by the F statistics.

Household income after taxes was introduced in the multiple regression model as a continuous variable.<sup>2</sup> Income had a statistically significant effect on the consumption of all nutrients with the exception of energy and the nutrients calcium, iron, thiamin, and riboflavin. The positive effect of income was statis-

tically significant at the 0.001 level for fat, niacin, and ascorbic acid; it was statistically significant at the 0.01 level for protein, and at the 0.05 level for vitamin A.

One factor which may have dampened the expected positive relationship between income and the consumption of some of the individual nutrients might be a negative relationship between purchasing efficiency of nutrients and income. Abdel-Ghany (1974) indicated that the quantities of nutrients that were obtained per dollar decreased as household income increased. In other words, at higher income levels, it would appear households substitute more expensive and less nutritious foods for cheaper and more nutritious foods. Kellerman (1972) also found no significant relationship between income and the daily recommended food serving adequacy.

AHMU was expected to be positively related to the con-

TABLE 3

Estimated Regression Coefficients and F Values for Selected Household Characteristics and Consumption of Energy (in Kilocalories)

Variable	Regression Coefficient	F
Household Income	210.414	1.456
Adjusted Household Meal Unit	954.066***	3469.401
Education of Female Homemaker (Years)	-787.603***	21.340
Age of Female Homemaker:		
14-24 years	-1172.821	0.584
25-44 years	omitted	
45-64 years	-4065.044***	13.595
65 and over	-9230.199***	31.688
Race of Female Homemaker:		
white	omitted	
black	-1547.392	1.413
Region:		
Northeast	-3156.018**	7.418
North Central	-2020.474	3.060
South	omitted	3.836
West	-2588.571	

$R^2 = 0.614***$

$F = 612.462$

\*\*Significant at .01 level

\*\*\*Significant at .001 level

<sup>1</sup>Since there is one degree of freedom to be associated with the explained sum of squares, then  $F = t^2$ . The test of significance in this case using F or t is identical. For further discussion see: (Snedecor and Cochran, 1956).

<sup>2</sup>The ordinally-scaled income variable in the data was treated as an interval variable in the analysis. Labovitz (1967) argues that the results obtained using ordinal data is rarely different from the results obtained using interval data.

sumption of individual nutrients, since its formulation depends mainly upon family size, composition, and number of meals eaten at home. Clearly, that was the case; the coefficients of AHMU were statistically significant at 0.001 level for all nutrients, and the relationship was positive for all nutrients.

Education of female homemaker had a negative effect on the consumption of all individual nutrients with the exception of ascorbic acid. The negative effect was statistically significant at the 0.001 level for protein

TABLE 4

Estimated Regression Coefficients and F Values for Selected  
Household Characteristics and Consumption of  
Protein (in Grams)

Variable	Regression Coefficient	F
Household Income	16.855**	8.410
Adjusted Household Meal Unit	30.173***	3124.728
Education of Female Homemaker (Years)	-18.826***	10.979
Age of Female Homemaker:		
14-24 years	-76.073	2.211
25-44 years	omitted	
45-64 years	-78.457**	4.560
65 and over	-271.926***	24.766
Race of Female Homemaker:		
white	omitted	
black	-79.357	3.348
Region:		
Northeast	31.799	0.678
North Central	51.765	1.809
South	omitted	
West	62.717	2.028

$R^2 = 0.592***$

F = 557.745

\*\*Significant at .01 level

\*\*\*Significant at .001 level

TABLE 5

Estimated Regression Coefficients and F Values for Selected  
Household Characteristics and Consumption of  
FAT (in Grams)

Variable	Regression Coefficient	F
Household Income	35.442***	15.249
Adjusted Household Meal Unit	42.316***	2520.140
Education of Female Homemaker (Years)	-34.640***	15.242
Age of Female Homemaker:		
14-24 years	-74.517	0.870
25-44 years	omitted	
45-64 years	-186.789**	10.599
65 and over	-520.608***	37.223
Race of Female Homemaker:		
white	omitted	
black	44.193	0.426
Region:		
Northeast	-182.632**	9.172
North Central	-59.627	0.984
South	omitted	
West	-135.173*	3.862

$R^2 = 0.554***$

F = 478.553

\*Significant at .05 level

\*\*Significant at .01 level

\*\*\*Significant at .001 level

TABLE 6

Estimated Regression Coefficients and F Values for Selected  
Household Characteristics and consumption of  
Calcium (in Milligrams)

Variable	Regression Coefficient	F
Household Income	-8.354	0.015
Adjusted Household Meal Unit	331.357***	2659.774
Education of Female Homemaker (Years)	-13.077	0.037
Age of Female Homemaker:		
14-24 years	-301.948	0.246
25-44 years	omitted	
45-64 years	-1957.607***	20.037
65 and over	-3019.995***	21.560
Race of Female Homemaker:		
white	omitted	
black	-3622.542**	49.234
Region:		
Northeast	-332.637	0.524
North Central	-478.769	1.092
South	omitted	
West	59.990	0.013

$R^2 = 0.547***$

F = 464.627

\*Significant at .05 level

\*\*Significant at .01 level

\*\*\*Significant at .001 level

TABLE 7

Estimated Regression Coefficients and F Values for Selected  
Household Characteristics for Consumption of  
Iron (in Milligrams)

Variable	Regression Coefficient	F
Household Income	-0.098	0.007
Adjusted Household Meal Unit	5.672***	2706.481
Education of Female Homemaker (Years)	-5.483***	22.829
Age of Female Homemaker:		
14-24 years	-2.081	0.041
25-44 years	omitted	
45-64 years	-17.309*	5.441
65 and over	-56.415***	26.132
Race of Female Homemaker:		
white	omitted	
black	-.466	0.003
Region:		
Northeast	-11.741	2.266
North Central	-3.177	0.167
South	omitted	
West	-3.485	0.153

$R^2 = 0.550***$

F = 470.490

\*Significant at .05 level

\*\*Significant at .01 level

\*\*\*Significant at .001 level



TABLE 8

Estimated Regression Coefficients and F Values for Selected  
Household Characteristics and Consumption of  
Vitamin A (in IU)

Variable	Regression Coefficient	F
Household Income	2509.583*	6.074
Adjusted Household Meal Unit	1797.409***	361.236
Education of Female Homemaker: (Years)	-648.907	0.425
Age of Female Homemaker:		
14-24 years	-12323.089	1.890
25-44 years	omitted	
45-64 years	2692.015	0.175
65 and over	-7627.607	0.635
Race of Female Homemaker:		
white	omitted	
black	40710.525***	28.701
Region:		
Northeast	23872.033***	12.451
North Central	891.804**	0.017
South	omitted	
West	19992.118**	6.712

$R^2 = 0.158***$

F = 72.414

\*Significant at .05 level

\*\*Significant at .01 level

\*\*\*Significant at .001 level

TABLE 9

Estimated Regression Coefficients and F Values for Selected  
Household Characteristics and Consumption of  
Thiamin (in Milligrams)

Variable	Regression Coefficient	F
Household Income	-0.103	1.112
Adjusted Household Meal Unit	0.510***	3150.460
Education of Female Homemaker (Years)	-0.462***	23.320
Age of Female Homemaker:		
14-24 years	0.554	0.413
25-44 years	omitted	
45-64 years	-1.555*	6.312
65 and over	-3.745***	16.549
Race of Female Homemaker:		
white	omitted	
black	0.878	1.444
Region:		
Northeast	-0.433	0.443
North Central	-0.184	0.085
South	omitted	
West	-0.348	0.220

$R^2 = 0.580***$

F = 531.419

\*Significant at .05 level

\*\*Significant at .01 level

\*\*\*Significant at .001 level

TABLE 10

Estimated Regression Coefficients and F Values for Selected  
Household Characteristics and Consumption of  
Riboflavin (in Milligrams)

Variable	Regression Coefficient	F
Household Income	-0.097	0.423
Adjusted Household Meal Unit	0.697***	2556.614
Education of Female Homemaker (Years)	-0.353*	5.912
Age of Female Homemaker:		
14-24 years	-1.389	1.129
25-44 years	omitted	
45-64 years	-3.017**	10.331
65 and over	-6.396***	20.990
Race of Female Homemaker:		
white	omitted	
black	-3.026**	7.458
Region:		
Northeast	2.205*	4.995
North Central	0.968	0.968
South	omitted	
West	1.428	1.609

$R^2 = 0.538***$

F = 448.020

\*Significant at .05 level

\*\*Significant at .01 level

\*\*\*Significant at .001 level

TABLE 11

Estimated Regression Coefficients and F Values for Selected  
Household Characteristics and Consumption of  
Riboflavin (in Milligrams)

Variable	Regression Coefficient	F
Household Income	5.597***	14.222
Adjusted Household Meal Unit	6.703***	2364.795
Education of Female Homemaker (Years)	-7.533***	26.960
Age of Female Homemaker:		
14-24 years	-33.515**	6.580
25-44 years	omitted	
45-64 years	-16.605	3.132
65 and over	-74.912***	28.823
Race of Female Homemaker:		
white	omitted	
black	-45.842***	17.131
Region:		
Northeast	14.900	2.283
North Central	19.719*	4.025
South	omitted	
West	8.926	0.630

$R^2 = 0.531***$

F = 436.697

\*Significant at .05 level

\*\*Significant at .01 level

\*\*\*Significant at .001 level

TABLE 12

Estimated Regression Coefficients and F Values for Selected Household Characteristics and Consumption of Ascorbic Acid (in Milligrams)

Variable	Regression Coefficient	F
Household Income	81.132***	82.168
Adjusted Household Meal Unit	22.728***	747.542
Education of Female Homemaker (Years)	18.311*	4.380
Age of Female Homemaker:		
14-24 years	-226.968**	8.298
25-44 years	omitted	
45-64 years	128.027*	5.120
65 and over	-30.476	0.131
Race of Female Homemaker:		
white	omitted	
black	-10.285	0.024
Region:		
Northeast	280.569***	22.259
North Central	-52.289	0.778
South	omitted	
West	36.689	0.293

$R^2 = 0.306***$

F = 169.798

\*Significant at .05 level

\*\*Significant at .01 level

\*\*\*Significant at .001 level

fat, iron, thiamin, and niacin as well as for energy; it was statistically significant at the 0.05 level for riboflavin. The positive effect for ascorbic acid was statistically significant at the 0.05 level.

The findings of this paper with respect to the negative relationship between the education of the female homemaker and the consumption of most of the individual nutrients are congruent with the findings of Madden and Yoder (1972) as well as Kellerman (1972) who found that education was not significantly related to adequacy of nutritive intakes and the daily recommended food servings adequacy in their respective studies. Also, Wilson and Lamb (1968) conducted a study of food beliefs among women and concluded that college graduates composed the largest group who accepted questionable beliefs about food. This negative relationship is perhaps due to the higher opportunity cost of time in the home among more highly educated women. The higher the value of the homemaker's time based on her earning opportunities outside the home, the higher the cost of preparing meals at home, and the fewer meals prepared. This might lead to either simple preparation of meals at home, including using foods that come into the home in highly semi-prepared form, or eating away from home more often. In either case, the attainment of a well balanced diet might be secondary to the minimization of time spent in food preparation.

The partial regression coefficients for the age of female homemaker set of dummy variables indicate that the 65-and-older category showed a statistically significant negative difference from the omitted category at the 0.001 level with respect to all nutrients with the exception of vitamin A and ascorbic acid. The 45-64 years of age category showed a statistically significant negative difference from the omitted group at the

0.001 level for energy and calcium; at the 0.01 level for protein, fat, and riboflavin; at the 0.05 level for iron, thiamin, and ascorbic acid. The 14-24 years of age group showed a statistically significant negative difference from the omitted category at the 0.01 level only for niacin and ascorbic acid.

Young, Waldner, and Berresford (1956) found a negative relationship between age and nutritional level. Jalso, Burns, and Rivers (1965) concluded that as age increased, valid nutritional opinions and practices decreased. Madden and Yoder (1972) found that age of the homemaker has a negative effect on the adequacy of all nutrients with the exception of iron. Kellerman (1972) also found a significant negative relationship between the age of homemaker and the daily recommended food serving adequacy

Black households consumed less calcium and niacin (significant at the 0.001 level) as well as riboflavin (significant at the 0.01 level) than did white households. Black households consumed more vitamin A than did white households (significant at the 0.001 level). There was no significant difference in the consumption of energy, protein, fat, iron, thiamin, and ascorbic acid between black and white households.

Regional variations in food consumption and nutrition may arise because of differences in customs, costs of living, availability of certain foods, and effects of climate. However, in this study relatively few differences were found between household characteristics and household consumption of individual nutrients.

The partial regression coefficients of the Northeast region were negative and statistically significantly different from the omitted region (i.e., the South) at

the 0.01 level for energy and fat; positive and statistically significantly different at the 0.001 level for vitamin A and ascorbic acid; and at the 0.05 level for riboflavin.

North Central showed to be positive and statistically significantly different from the omitted region at the 0.05 level for only niacin.

The Western region was negatively and statistically significantly different at the 0.05 level for fat; positively and statistically significantly different at the 0.01 level for vitamin A.

#### SOME IMPLICATIONS

Clearly, there is need for better understanding of nutrition and of the need for food consumption patterns which better fulfill nutritional needs. This prevails regardless of education or income and, for the most part, age. Therefore, it would appear there is a very present need for education, broadly perceived, both formal and particularly informal (e.g., extension programs, etc.) aimed at adults as well as younger people and probably in various educational modes. As part of this, possibilities may exist for imaginative use of short spots dealing with nutrition concepts for TV or radio airing. Other communication techniques undoubtedly have pertinence, too.

Perhaps the food processing industry needs to continue to examine its role in supplying the food and nutritional needs of the population. It may be that fortification, always in consultation with the best nutritional advice, needs to be re-examined, modified and carried further in some instances; particularly in foods commonly consumed by older people may this need exist. It may be especially important for the food industry to examine its advertising practices and to adopt a code aimed at providing information really useful to consumers in making choices affecting their nutritional status.

In any event, barring substantial shifts in the underlying relationships, it would appear that increases in general educational levels or an upward drift in incomes can't be assumed to bring an automatic improvement in household nutritional status.

#### REFERENCES

- Abdel-Ghany, M. "Influences of U. S. Urban Household Characteristics on Nutritive Intakes and Quality of Food Consumption," Ph.D. dissertation, University of Missouri, Columbia, 1974.
- Jalso, S. B., M. M. Burns, and J. M. Rivers. "Nutritional Beliefs and Practices," Journal of the American Dietetic Association 47 (October 1965), 263-68.
- Kellerman, B. J. "Correlates of Food Shopping Efficiency, Dietary Adequacy, Food Variety, and Food Expenditures Among Low-Income Rural Missouri Households," Ph.D. dissertation, University of Missouri, Columbia, 1972.
- Labovitz, S. "Some Observations on Measurement and Statistics," Social Forces 46 (December 1967), 151-160.
- Madden, J. P. and Marion Yoder. Program Evaluation: Food Stamps and Commodity Distribution in Rural Areas of Central Pennsylvania. Agricultural Experiment Station Research Bulletin 780. University Park, Pa.: The Pennsylvania State University, 1972.
- Snedecor, G. W. and William Cochran. Statistical Methods. 5th ed. Ames, Iowa: The Iowa State University Press, 1956.
- U. S. Department of Agriculture. "Cost of Food at Home," Family Economics Review (December 1965), 5.
- U. S. Department of Agriculture. Food Consumption of Households in the United States, Spring 1965. Household Food Consumption Reports, Nos. 1-10, Washington, D.C., 1968.
- Wilson, M. and M. Lamb. "Food Beliefs as Related to Ecological Factors in Women," Journal of Home Economics 60 (February 1968), 115-28.
- Young, C. M., B. G. Waldner, and K. Berresford. "What the Homemaker Knows About Nutrition," Journal of the American Dietetic Association 32 (March 1956), 218-22.

# THE FOOD PROBLEM AND INCOME ADEQUACY<sup>1</sup>

Flora Williams, Purdue University

## Abstract

"Having enough food" was examined among 1275 urban families for relationships with income measures, family types, financial problems, and social processes. Ten percent "often" had this problem and 58 percent "never" had it. Of those with income comparable to food stamp eligibility, 42 percent "often" or "sometimes" did not have enough food. Associated were income measures, financial problems, ethnicity, years of family formation, husband's and oldest child's ages, husband's occupation, single parent status, SES, and who handles money.

## The Food Problem and Income Adequacy

The current economic condition of recession and greatly increased prices has brought attention to the most critical of consumer products - our daily food. Currently there is national discussion over the income measure used for determining food stamp eligibility, benefits, and poverty line. The official government poverty line and food stamp benefits are based on the cost of a minimal diet assuming one third of the income is spent for food. The availability of food stamps for eligible consumers has been required since 1974. The problem of enough food is considered the first to solve in the poverty cycle in order that consumers can continue to live, learn, produce, earn, and overcome other barriers to obtain well-being.

Many assumptions have been accepted in establishing the specific food plan for the basis of the poverty line and the maximum income and assets used for determining eligibility for food stamps. The poverty line and maximum income have been taken as "givens" but this study addresses these measures directly. The frequency of the food problem for consumers under the poverty line and those eligible for food stamps needs to be examined for further questioning those assumptions. How frequently do consumers have the problem of not enough food until there is money to buy more? Are there factors related to the food problem other than income; and which measure of income would be most appropriate for analyzing the food problem? What types of consumers have the food problem and, if they do, what other financial problems are also likely to be present?

Most households' eligibility is based on maximum levels of money income. If the asset limitations (\$1500 of countable resources of all members if under age 60) were used on the basis for eligibility, perhaps 15 or 20 percent additional households would be disqualified. "The assets maximums apply to households of every size, but net income maximums increase with household size to register rising household needs" (MacDonald, 1975, p.4).

Income maximums for food stamp eligibility are based on household net income, which is total money income less ten percent from wages or salaries not to exceed \$30 a month, mandatory deductions from earnings, medical expenses exceeding \$10 a month, child care payments if permitting a member to be employed, tuition and mandatory education fees, and "excess shelter costs" exceeding 30 percent of the net income remaining after other deductions.

Most AFDC recipients without earnings would be eligible as would, also, those with earnings when deductions are used to arrive at net income. A working father with three dependents, filing a joint income tax return with a \$7200 gross earnings would have a net income less than the \$6000 which was the annual maximum for a four-person household just by federal tax and Social Security withholdings being deducted (Bleckman, et al, 1974 p. 192). (This figure is comparable to the disposable money income figure used in this study.) The shelter deduction, in addition, could cause total money income to exceed net income by a substantial margin (Bickel and MacDonald, January 1975, p. 4).

The net income is considered to be a better indication of household well-being than gross income. Food stamp benefits, as well as eligibility, are determined by the net income instead of gross income so that special circumstances reflected in the deductions are considered (MacDonald, 1975, p.6).

MacDonald (1975, p.6) discussed the issue of which income-assets measure to use for eligibility: "Yet however desirable the extensive list of allowable deductions may be from an equity standpoint, a substantial portion of total program administrative costs is directly attributable to attendant complications for the benefit and eligibility determination process." He noted that considerable time is devoted to the calculation of net income and there is likelihood of payment errors when complicated calculations are necessary. Bickel and MacDonald (January 1975, p.5) report that there is no readily usable data for comparing the difference between census gross income and food stamp net income of low-income households.

Congressional proposals may be enacted which would revise the net income definition to incorporate a standard deduction from gross income. MacDonald found in his study (1975, p.7) "that the food stamp net income eligibility standards are quite comparable to the gross income levels for 125 percent of the non-farm poverty line." In Projector and Weiss (1966, Table 39), the income level of \$6400 per family of four which was 125 percent of the poverty line in 1974 dollars closely compared to the current food stamp net income standard of \$6000 per four-person household (Bickel and MacDonald, January 1975, p. 35).

Eligibility for food stamps is based on current monthly incomes because of the fluctuating incomes among many low income households (which yields an estimated number actually higher than if annual income data were used). However, at higher income levels, income fluctuations are also prevalent and at the lowest incomes, many incomes are stable.

Food is only one of many pressing needs of low-income consumers. In addition to the food problem, the government has tried to alleviate several other financial problems such as housing and medical care. Consumers have a choice in the management of their resources to

<sup>1</sup>This study contributes to inter-regional Agricultural Experiment Station research project NC-90, "Factors Affecting Patterns of Living of Disadvantaged Families." Cooperating states: California, Hawaii, Illinois, Indiana, Iowa, Kansas, Missouri, Nebraska, Nevada, Ohio, Texas, Vermont, Wisconsin. Journal Paper No. 6002, Purdue Agricultural Experiment Station, West Lafayette, Indiana, 47907.

allocate income to those considered most pressing. MacDonald (1975, p.12) stated "the food stamp program could do eligibles no good if they chose not to earmark a sizable proportion of their budget for food because other, equally pressing needs might arise before they next received cash."

The objectives of this study were to

- determine the relationship of the food problem (not enough food) to income measures (objective and perceived), income dependability, and income source.
- compare relationships of the food problem with income index, disposable income, and net income, and perceived adequacy of income.
- identify the relative order of the food problem with other financial problems assuming the order depends on reported frequency based on which expenditures seem more pressing or have legal consequences in relation to others.
- determine the relationship of the food problem to other problems and to fixed commitments.
- determine the relationship of the food problem to demographic characteristics of consumers.
- examine the relationship between selected family processes and the food problem. Selected processes included: who decides the use of money, who handles the money, and the frequency of eating together as a family.

One hypothesis was that below a low income level, the food problem would not be related to variables other than objective measures of money income, but at higher income levels, other variables representing management and priorities would be related. Two-parent families within similar income categories were hypothesized to have more food problems than single-parent families, especially if most of the single parents received welfare payments and more stable incomes. Nonemployed wives would report less problems, it was hypothesized, because they might have more time to do careful shopping and food preparation than employed wives and nonemployed female heads of families. Language spoken, if other than English, was hypothesized to be a barrier to the absence of the food problem. The higher educated consumers and older consumers were hypothesized to have the food problem less frequently than others because of their abilities and experience. The income index, an objective measure of income adequacy, used for establishing poverty status, was hypothesized to be more related to the food problem than disposable income, net income, perceived adequacy of income, or income dependability because it was believed to be a more accurate measure of income adequacy.

#### Methodology

An interdisciplinary, regional research project, NC-90, entitled "Factors Affecting Patterns of Living in Disadvantaged Families" provided the data used for this study. The disciplines participating in the project were: rural and family sociology, child and educational psychology, economics, family economics, home management and child development. Thirteen state experiment stations were represented in the project. Members of the NC-90 committee developed the interview schedule and provided instructions in training interviews. Interviewers were indigenous to the six urban disadvantaged areas supplying data for this part of the study. Data were collected during 1970 and subsequently analyzed. Sufficient detail was obtained to calculate net income identical to that used for determining food stamp eligibility. Other than these data, there were no readily usable data in the country for comparing net income with other income measures.

#### Source of the Data

Data collected randomly from families in six metropolitan areas were analyzed for purposes of this study. The

homemaker served as the respondent. Area samples were drawn by the Survey Section of the Iowa State University Statistical Laboratory for the studies in Indiana, Ohio, and Nevada. East Chicago, Indiana and Toledo, Ohio, families were selected from areas designated as poverty tracts by the U.S. Department of Commerce, Bureau of the Census (1966). The population from which the Nevada sample was drawn consisted of families living in designated "low cost housing" areas as determined through the compilation of information by the City Planning Department of Las Vegas. Hawaii data were collected within eight census tracts in metropolitan Honolulu on the island of Oahu. In these tracts during the three year period of 1964-67, 40 percent or more of the families had an annual income of \$3,000 or less. The Statistical Laboratory at the University of Illinois sampled a low income area based upon the assessed valuation of property in Champaign-Urbana, Illinois. In Superior, Wisconsin the families were selected from an economically depressed area.

To be eligible for the personal interview, a family had to have at least one child under age 18 and a homemaker, either gainfully employed or not, between 18 and 64 years of age. Families varied widely in the distribution of income, education, number in family, ethnicity and frequency of the food problem.

#### Characteristics of the Sample

Although areas were considered disadvantaged, the majority of the families would not be classified as low income as defined by the Social Security Administration (Orshansky, 1969). However, the great majority would be classified as lower socio-economic classes. Characteristics of the families in this sample from metropolitan areas varied widely in other demographic characteristics. By ethnicity, 46 percent were white, 34 percent were black, 5 percent were oriental, 12 percent were Spanish-American, and 9 percent were of other ethnic origins. The mean total disposable income was \$8,960.47. Twenty-two percent of the families in the sample had over \$10,000 income. As defined by the Social Security Administration, 15 percent of the families were classified as below the poverty level (below 100 income index) and about ten percent were classified in the near-poverty level (between 100 and 125 income index). The mean number in the family was five. The mean age of the husband was 37. Median years of schooling completed for husband and wives was 10.8.

#### Data Analysis

Data from these six metropolitan areas were combined for analysis purposes. The number of families supplying usable data was 202 from Hawaii, 287 from Illinois, 193 from Indiana, 225 from Nevada, 170 from Ohio, and 208 from Wisconsin. In some analyses the number of cases was less when information was incomplete.

Initially, fine levels of income index were observed for determining any pattern of acute change in the food problem as income index increased. To control for the effect of income, additional analyses were made of families in various income groups.

Statistical tests for determining the associations and differences between frequency of the food problem and the selected variables were chi-square analysis and analysis of variance. The latter was used with data involving the food problem and income measures. The Pearson's product-moment correlation was used to examine the association of the score of the food problem frequency and other variables. The .05 level of significance was used to determine statistically significant differences and relationships.

Gamma, a measure of association for ordinal scales, was used to examine the association between frequency of the food problem and the frequency of each of the other financial problems. A gamma value of .4 would indicate that the predictability of one factor would be improved by 40 percent by observing another factor. A value of .35 for gamma was judged to be a significant association for this study.

#### Selected Variable Definitions

The food problem was one of 11 financial problem areas selected by the NC-90 committee after their review of literature and conjecture as those believed to be meaningful to families. The general question asked in the schedule was: "Aside from not having enough money, which of the following problems do you have and how often do you have this problem?" The specific question referred to "not enough food until there is money to buy more." "How often" was checked on the basis of the frequency never, seldom, sometimes, and often which were assigned scores of 1, 2, 3, and 4 respectively, and treated as linear.

The food problem was examined in relation to the families' income index. The index is a measure of income adequacy based on the cost of food for a family considering its size, sex and age composition (Orshansky, 1969, p. 38). The cost of food was calculated from food plans developed by the United States Department of Agriculture. An income index of 100 represented money income of a family three times the cost of the "economy food plan." The amount of money allocated to food in the economy plan was "the minimum that the Agriculture Department said could still provide American families with an adequate diet." Less than one of ten families spending that amount or less were expected to have had good diets (Peterkin and Clark 1969, p. 8). A family with an income of less than or equal to 100 was classified as poor. An income index of 125 represented money income of a family three times the cost of the "low-cost food plan" and therefore, families with income indices of 100 to 125 were classified as near-poor. About three of ten families spending that amount were expected to have good diets. The income index (Orshansky, 1969) seemed to the researchers to be the most refined and reasonable measure of income adequacy to date that could be used in assessing the family's economic well-being.

Money income was the flow of purchasing power received by the family during the previous 12 months from economic activities (earned income), investment, social security, job related benefits, armed services insurance, welfare payments, and legal arrangements. Total disposable income consisted of total money income plus pay check deductions minus state and federal income taxes and Social Security contributions.

Net income was money income less ten percent from wages or salaries not to exceed \$30 a month, mandatory deductions from earnings, medical expenses exceeding \$10 a month, child care payments if permitting a member to be employed, tuition and mandatory education fees, and "excess shelter costs" exceeding 30 percent of the net income remaining after other deductions.

Perceived adequacy of income was assessed by asking the following question of the homemakers: "To what extent do you think your income is enough to live on?" Responses were: 1) not at all adequate, 2) can meet necessities only, 3) can afford some of the things we want but not all we want, 4) can afford about everything we want, and 5) can afford about everything we want and still save money.

Fluctuations in income have been found to play an important role in consumption behavior (Manning, 1960) but

stronger evidence of that role is needed. Income dependability referred to the regularity with which income was received. The homemakers were questioned as to the dependability of income and their responses were compared with the actual receipt of income. The dependability of family income was classified as: 1) "not dependable" at all, if it never or seldom knew more than a month in advance how much it would have and when it would get any money income; 2) "fluctuating," if income varied in timing or amount or both; and 3) "steady," if income was dependable and stable.

Sources of income were categorized into 1) social security only, 2) welfare plus earnings, 3) welfare only, and 4) earnings and no welfare. Welfare sources included aid to the blind, aid to permanently and totally disabled, old age assistance, aid to families with dependent children, general assistance, and private agency. Earnings included sources of salary and wages, profits from business, bonuses and commissions, job related benefits, armed service benefits plus money from legal arrangements, gifts, and inheritances.

The percentage of income allocated to financial commitments were calculated from the "bills or expenses" the family was "supposed to meet regularly" including things they felt were "rather fixed, that they were obliged to or had promised to pay every week or month, or that were taken out of a paycheck."

Initially, languages were English, Spanish, French, Japanese, English plus another language, and others. For purposes of analysis, English, Spanish, English plus other, and other languages were compared.

The number of years the family had been formed was an indicator of the life cycle. The interviewer asked each respondent when the family was started, and then probed whether the number of years marked the beginning of married life, the birth of the first child, or when the oldest child came to live with the family. It was hypothesized that the longer the family had been established, the more experience it would have with handling money and managing the food problem.

Occupational levels of husbands and wives were initially classified 1) higher executive and professional, 2) business manager, 3) administrative personnel, 4) clerical and sales, 5) skilled, 6) semi-skilled, and 7) unskilled. However, classifications one through three were combined for purposes of analysis to meet criteria for cell size.

Employment status of women was analyzed by comparing female heads and wives who were employed and who were non-employed in these four classifications.

The respondent's education was reported on the questionnaire. Actual years of schooling were grouped into the usual classifications for analysis.

Family type was analyzed by the simple classification of two-parent family and one-parent family. In all cases the one-parent family had a female head.

Rather than relying on an existing scale, a question was developed which asked for the homemaker's perception of who (wife, husband, or husband and wife together) mainly decided "how the money is used." Another question was asked in order to identify the effective agent in the family who actually implements decisions. The item asked for the homemaker's perception of who (wife, husband, or husband and wife together) mainly handles money matters. Husband's presence during the wife's responses to these questions was viewed as a necessary preliminary measure.

Findings

Data from the six areas were pooled for purposes of analysis. All areas were designated disadvantaged, located in central cities, and were sampled randomly. All together, 10 percent of the families often had the food problem and 58 percent never had the food problem. Even though the frequency of the food problem differed among states as shown in Table 1, most of the other variables selected for analysis were associated with the food problem when all the families were examined. Furthermore, for families with income index under 100 (poverty) or under 125 (near-poverty) the frequency of the food problem did not differ by the six areas.

TABLE 1

Frequency of the problem "not enough food" for families by state sample.

State	Frequency of food problem (percent)				Total number
	Never	Seldom	Some times	Often	
Hawaii	60	11	23	6	202
Illinois	68	13	14	5	287
Indiana	42	16	17	26	193
Nevada	48	21	20	12	224
Ohio	55	12	22	10	170
Wisconsin	71	11	12	7	198
Total percent	58	14	18	10	

Measures of Income

Income index, an objective measure of income adequacy listed in Table 2, was associated with the frequency of the food problem ( $r = -.25, p < .01$ ). For all families, income index increased as the frequency of the food problem decreased. This significant relationship remained in three of the six disposable income levels, less than \$4000, \$4000 through \$5999, and \$6000 through \$7999. As shown in Table 3, some families often had the food problem at every index level, but a pattern is observed and a shift noted at the 150 income index bracket. Up until then "never" was reported less than would be expected and "seldom", "sometimes", and "often" were reported more than would be expected. Almost half of the families under 100 index and one-third of the families from 100 to 125 index "sometimes" or "often" had the food problem. Of those families with income comparable to that for food stamp eligibility, 42 percent "often" or "sometimes" did not have enough food. Analyzing the families another way as shown in Table 4, income index for those "never" having the problem was 205.66 while for those "often" was 143.41 and "sometimes" was 125.42.

TABLE 2

Pearson's product-moment correlations of food problem and income measures at various disposable income levels.

Disposable income	Number	Income measures			
		Income index	Disposable income	Net income	Perceived adequacy
All families	1275	-.25**	-.28**	-.25**	-.34**
Less than \$4000	252	-.17**	-.14*	-.08	-.25**
\$4000-\$5999	267	-.25**	-.07	-.03	-.27**
\$6000-\$7999	263	-.12*	.06	.03	-.29**
\$8000-\$9999	180	.04	.06	.05	-.25**
\$10,000-\$11,999	121	-.07	-.10	.09	-.25**
\$12,000-above	192	-.01	-.04	-.25**	-.31**

\*Significant at the five percent level.  
\*\*Significant at the one percent level.

TABLE 3

Frequency of the food problem among all families at different income levels.

Income level	Total number	Frequency of the food problem (%)			
		Never	Seldom	Sometimes	Often
Income index					
Under 74	144	35	17	35	13
75 - 99	134	41	15	25	19
100-124	135	52	16	23	10
125-149	144	50	19	18	13
150-174	113	63	16	15	6
175-199	121	61	12	15	12
200-224	119	72	12	10	6
225-249	72	71	13	7	10
250-274	52	75	12	12	2
275-299	42	81	10	0	10
300-above	166	80	7	8	5
Total	1242	58	14	18	10

TABLE 4

Anova of income measures by frequency of the food problem.

Food problem	Income index**	Perceived adequacy**	Disposable income**	Net income**	Number
Never	205.66	3.23	\$8543	\$7555	712
Seldom	158.72 <sup>a</sup>	2.67 <sup>a</sup>	6347 <sup>a</sup>	5631 <sup>a</sup>	173
Some-times	125.42 <sup>a</sup>	2.45 <sup>a</sup>	5663 <sup>a</sup>	4995 <sup>a</sup>	220
Often	143.41 <sup>a</sup>	2.53 <sup>a</sup>	5701 <sup>a</sup>	5033 <sup>a</sup>	128
					1233
Range	22 thru 1667	1 thru 5			

\* Significant at the five percent level.

\*\* Significant at the one percent level.

<sup>a</sup> Neuman-Keuls sequential range test indicated no differences between means with the same superscript.

Disposable income was also associated with the frequency of the food problem ( $r = -.28, p < .0$ ) for all families and those with income under \$4000 ( $p < .05$ ) (Table 2). However, by chi square analysis summarized in Table 5, disposable income was not significantly associated for families with income index under 125 and over 200. It was associated between the ten and five percent levels. The shift in the pattern of those never having the problem occurred at the \$6000 bracket. About 60 percent of the families below \$2000 and about half of the families with incomes from \$2000 and \$3000 had the food problem "sometimes" or "often" as shown in Table 6. Between \$4000 and \$5000, 36 percent had the problem "sometimes" or "often". A third had the problem with this frequency when incomes were between \$5000 and \$6000 and one-fourth had it with this frequency when incomes were between \$6000 and \$7000. The mean disposable income for families who "often" had the problem was \$5701 and for families who "never" had the problem was \$8543. (Table 4).

Net income, the measure currently used for determining food stamp eligibility, was associated with the frequency of the food problem ( $r = -.25, p < .01$ ) in the analysis for all families and for those with disposable incomes of \$12,000 and above (Table 2). As a statistical measure net income did not explain variation in the food problem any better than other income measures. Results as shown in Tables 2 and 4 indicate net income was comparable to income index and disposable income as a statistical measure. The mean net income for families who "often" had the food problem was \$5033 and for families who "never" had the problem was \$7555 (Table 4). As observed in Table 6, the shift in the pattern of the occurrence of the food problem appeared at the \$5000 bracket.

Perceived adequacy of income was associated with the food problem ( $r = -.34, p < .01$ ) and this relationship remained



in all the income index groups and most of the disposable income levels as observed in Table 2. Over half of the families who said their income was not adequate had the food problem "often" or "sometimes"; but 13 percent who said they could afford about everything had the problem "often" or "sometimes" (Table 6). The average perceived adequacy for those families who had the problem "often" was 2.53, between the responses of "could afford necessities only" and "could afford some luxuries!"

Income dependability was associated with the food problem at the .001 level of significance for all families and those under 100 income index but at the .06 level and .07 level for families under 125 index and over 200 index, respectively (Table 5). Of all families with income assessed as "not dependable", 57 percent "often" or "sometimes" had the problem, whereas of those with steady income, 28 percent had it this frequently as figured from Table 6.

TABLE 5

Significance level of chi-square obtained from analyses of income index groups by selected variables.

Selected variables	Income index groups			
	All families (n=1274)	Under 100 (n=311)	Under 125 (n=446)	Over 200 (n=451)
Income index	.00	--	--	--
Disposable income	.00	NS <sup>a</sup>	.09	.07
Perceived income adequacy	.00	.00	.00	.00
Income dependability	.00	.00	.06	.07
Source of income	.00	.00	.00	NS
Net income	.00	NS	.04	NS
Money problems of				
Rent or house payment	.00	.00	.00	.00
Kid's things	.00	.00	.00	.00
Doctor and medicine	.00	NS	.00	.00
Utilities	.00	.00	.00	.00
Large bills	.00	.00	.00	.00
Equipment maintenance	.00	.00	.00	.00
Clothing	.00	.00	.00	.00
Saving	.00	NS	.05	.00
Percent of income committed to credit	.00	NS	NS	.09
Ethnic status	.00	NS	.03	.00
Language spoken	.00	NS	NS	.00
Number of years family has been formed	NS	NS	.04	NS
Age of homemaker	.06	NS	NS	NS
Age of husband	.02	NS	.03	.00
Age of oldest child	NS	NS	.04	NS
Occupation of homemaker	.00	NS	NS	.06
Occupation of husband	.00	NS	.05	.00
Employment status of homemaker	.00	NS	.07	.00
Education of homemaker	.00	NS	NS	.05
Education of husband	NS	NS	NS	NS
Family type	.00	.08	.03	NS
Socio-economic status	.00	.00	.00	.00
Who decides use of money	NS	NS	NS	.06

Table 5 (Cont'd)	All gps	Under 100	Under 125	Over 200
Who handles money	.05	NS	.02	NS
Eating together	NS	NS	NS	NS

<sup>a</sup>Chi-squares were not significant at the ten percent level or beyond.

The frequency of the food problem was associated with the source of income except for families who had over 200 income index ( $p < .01$ ) (Table 5). Of those families with only social security or welfare income, about two-thirds had the problem "sometimes" or "often" whereas almost one-half of those with welfare payments and earnings and one-fifth of those with earning but no welfare payments had the problems "sometimes" or "often" (Table 6).

Other Problems and Fixed Commitments

Other financial problems listed in Table 7, were found associated with the frequency of the food problem ( $p < .01$ ) by chi square analysis. However, for families with income index under 100, the problems of doctor-medicine and savings were not found associated with frequency of food problem (Table 5). About one-half of the families who "often" had the food problem also "often" had the problems of kids' things, doctor or medicine, clothing, and savings; and about one-third also "often" had the problems of large bills and maintaining equipment. The range of gamma values indicated that when the food problem was observed, the prediction of the saving problem could be improved by 41 percent and of problems in kids' things, utilities, and clothing by 57 percent.

TABLE 6

Frequency of the food problem among all families at different income levels and results of chi-square test.

Income level	Frequency <sup>a</sup> of food problems (percent)			
	Number	Never	Seldom	Some-times Often
<u>Disposable income</u> (Chi-square 127.15*)				
0-1999	64	23	17	39 20
2000-2999	94	36	18	30 16
3000-3999	94	40	19	26 15
4000-4999	120	44	20	24 12
5000-5999	147	52	16	22 11
6000-6999	146	66	11	9 14
7000-above	610	70	11	12 7
<u>Net income</u> (Chi-square 107.02**)				
0-1999	125	38	14	30 18
2000-2999	102	36	22	28 14
3000-3999	119	44	17	24 16
4000-4999	168	49	19	22 10
5000-5999	150	62	11	17 11
6000-6999	123	65	11	13 11
7000-above	488	71	11	11 7
<u>Perceived income adequacy</u> (Chi-square 201.71**)				
not adequate	81	22	15	36 27
necessities only	233	35	20	34 11
some luxuries	760	61	15	14 10
about everything	82	82	6	9 4
everything & save	118	93	3	3 2
<u>Income dependability</u> (Chi-square 25.63**)				
not dependable	44	32	11	41 16
fluctuating	699	57	16	17 10
steady	529	61	12	18 10
<u>Sources of income</u> (Chi-square 133.63**)				
social security	9	22	11	44 22
welfare only	67	22	13	45 19
welfare & earnings	223	35	20	31 14
earnings-no welfare	921	66	13	12 9
other	54	67	9	17 7
Total		58	14	18 10

\* Significant at the five percent level.

\*\* Significant at the one percent level.

<sup>a</sup>Percentage frequencies do not total 100 due to rounding.

TABLE 7

Frequency of the food problem among all families by frequency of other problems and results of chi-square test.

Other problems	Frequency of food problems (percent)				
	Number	Never	Seldom	Some-times	Often
Rent or house payment ( $X^2_a = 163.77^{**}$ ; $\gamma = .43$ )					
Never	952	65	11	15	9
Often	20	10	15	25	50
Kids' things ( $X^2_a = 300.09^{**}$ ; $\gamma = .57$ )					
Never	372	84	6	7	4
Often	253	23	20	29	28
Doctor & medicine ( $X^2_a = 256.52^{**}$ ; $\gamma = .48$ )					
Never	726	71	9	14	6
Often	158	18	17	27	37
Utilities ( $X^2_a = 229.99^{**}$ ; $\gamma = .57$ )					
Never	999	67	10	14	9
Often	25	4	8	40	48
Large bills ( $X^2_a = 310.44^{**}$ ; $\gamma = .53$ )					
Never	685	74	36	34	6
Often	114	11	14	34	41
Equipment maintenance ( $X^2_a = 306.50^{**}$ ; $\gamma = .54$ )					
Never	776	72	12	12	5
Often	90	14	10	30	46
Clothing ( $X^2_a = 363.53^{**}$ ; $\gamma = .57$ )					
Never	546	80	8	8	4
Often	194	23	17	24	37
Saving ( $X^2_a = 132.76^{**}$ ; $\gamma = .41$ )					
Never	387	77	9	10	5
Often	479	40	18	26	16
Total number	1273				

<sup>a</sup>The frequencies of problems in all responses (never, seldom, sometimes, and often) were compared in the chi-square test.

<sup>\*\*</sup>Significant at the one percent level.

How does food expenditures rank in relation to other pressing needs for the allocation of income? If the assumption is held that the frequency of problems reported is inversely related to their perceived importance due to values or legal consequences, the following order might indicate the rank of pressing needs; utilities, rent or house payments, maintaining equipment, food, large bills, doctor and medicine, clothing and personal things, kids' things, and saving. That is, money may have been allocated to utilities first before other needs, and therefore, not allowed to become a problem. For these families, savings was reported by the greatest percentage of families because the other expenditures were taken care of before it.

Families with the greater percentage of income committed to fixed expenditures had a greater percentage reporting not enough food ( $p < .01$ ) as observed in Table 8. Of families with less than 20 percent of disposable income committed to fixed expenditures, 70 percent "never" had the food problem. Of those with 80 percent or more committed, 42 percent "never" had it, and 41 percent "often" had it. Consider Orshansky's assumption that food is one-third of the budget for low-income families and add other fixed costs such as housing. Then the finding that the majority of families also had 20 percent or more of income committed to credit is revealing. Although the association of the food problem and fixed commitment was significant in general, it was not significant for families under 100 or under 125. Evidently, fixed commitments as a percentage of income and the food problem was highly associated for families between 125 and 200 income index.

#### Demographic characteristics

Most of the demographic characteristics selected for analysis were associated with the food problem at the .01 level (Table 5). Significant ones for all families

TABLE 8

Frequency of the food problem among all families by frequency of financial commitments as percentage of total family income.

Percent of income committed	Frequency of food problems (%)				
	Number	Never	Seldom	Some-times	Often
Chi-square <sup>a</sup> = 73.83 <sup>**</sup>					
Less than 20 percent	107	70	8	16	6
20 - 29	264	67	13	13	6
30 - 39	262	65	14	14	7
40 - 49	219	56	14	18	12
50 - 59	146	54	15	18	13
60 - 69	79	43	14	32	11
70 - 79	37	35	11	30	24
80 and above	160	42	18	23	18
Total	1274	58	14	18	10

<sup>\*\*</sup>Significant at the one percent level.

<sup>a</sup> See footnote in Table 7.

are described in Table 9. Not associated were number of years family had been formed, age of oldest child, and husband's education. However, for families with income index under 125 the variables associated were ethnic status, number of years family had been formed, husband's age, husband's occupation, age of oldest child, employment status of homemaker ( $p < .07$ ), family type, and socio-economic status. However, for families under 100 income index none of these were significant except family type which came close ( $p < .08$ ). It was for families over 200 income index that ethnic status, language spoken, homemaker's occupation, homemaker's employment status, husband's occupation, and homemaker's education were significant.

A greater percentage of Blacks and Spanish-Americans had the food problem than whites. Families of the youngest and the oldest husbands had a greater percentage of reporting the food problem than others. Homemakers in the professional and unskilled occupations had a greater percentage reporting the food problem than other employed homemakers. But this pattern was reversed for husband's occupation where a greater percentage of the skilled and semi-skilled husbands had the problem than others. Although a greater percentage of female heads, regardless of being employed or not, reported the food problem more than others, this relationship did not hold for homemakers with family income index under 100. A smaller percentage of all homemakers with eighth grade education or less reported the food problem than those with education beyond high school but this relationship did not hold for families under 100 and 125 income index. A greater percentage of one-parent families had the food problem than two-parent families. Of upper-middle class families, 74 percent reported never having the problem compared to 46 percent of the lowest class.

#### Social Processes

How frequently the family ate together was not associated with the frequency of the food problem. Altogether 83 percent said they often ate together, 10 percent said they sometimes did, six percent said they seldom did, and one percent said they never did.

Who decides the use of money was not associated with the money problem but who carries out the decisions or handles the money was for all families and for those under 125 income index. When both husband and wife handled the money, a greater percentage (12%) reported "often" than when either the wife alone or the husband alone (8%) handled the money as shown in Table 10. Altogether, 68 percent said both husband and wife decided the use of money, 17 percent said the husband did, and 15 percent said the wife did. In handling the money, 46 percent of the wives were reported as mainly responsible,

30 percent of both husbands and wives, and 24 percent of the husbands.

Summary and Conclusions

TABLE 9

Frequency of the food problem among all families by significant demographic characteristics and results of chi-square.

Characteristics	Number	Frequency of food problems (%)			
		Never	Seldom	Some-times	Often
Ethnic status (Chi-square <sup>a</sup> = 73.66**)					
White	582	67	13	14	6
Black	432	47	16	20	17
Spanish-American	73	45	12	25	18
Others	171	57	12	28	3
Occupation of homemaker (Chi-square <sup>a</sup> = 44.69**)					
Prof., mgrs	57	52	31	7	10
Clerical, sales	189	68	13	11	8
Skilled	9	67	11	22	0
Semi-skilled	47	55	13	28	4
Unskilled	263	49	16	24	11
Family type (Chi-square <sup>a</sup> = 48.42**)					
Two-parent	874	64	12	14	9
One-parent	400	44	18	25	13
Age of husband (Chi-square <sup>a</sup> = 79.22**)					
Under 18	349	42	18	28	14
18-24	101	51	19	23	8
25-34	323	64	13	13	10
35-44	224	64	12	14	10
45-54	203	73	11	9	7
55 and over	74	60	8	22	11
Employment status of homemaker (Chi-square <sup>a</sup> = 65.22**)					
Employed, fem. head	225	46	20	22	12
Employed, wife	341	67	11	15	8
Unemployed, fe. hd	175	42	13	29	16
Unemployed wife	533	63	13	14	10
Education of homemaker					
Less than eight	510	63	14	14	9
9-11	407	51	17	20	13
12	156	78	9	8	5
13-17	192	43	13	30	14
Occupation of husband (Chi-square <sup>a</sup> = 67.05**)					
Prof., mgrs	105	82	6	7	6
Clerical, sales	178	69	12	11	9
Skilled	194	60	13	17	10
Semi-skilled	268	51	19	19	10
Unskilled	145	83	6	4	7
Socio-economic status (Chi-square <sup>a</sup> = 73.36**)					
Upper-middle	119	74	4	11	10
Middle	145	80	7	8	6
Lower middle	174	70	10	12	9
Upper-lower	242	62	18	12	9
Lower-lower	192	46	17	25	13
percent	100	58	14	18	10

<sup>a</sup>The frequencies of problems in all responses (never, seldom, sometimes, and often) were compared in the chi-square test.

\*Significant at the five percent level.

\*\*Significant at the one percent level.

TABLE 10

Frequency of the food problem among all families by significant social processes and results of chi-square.

Processes	Number	Frequency of food problems (%)			
		Never	Seldom	Some-times	Often
Who handles money (Chi-square <sup>a</sup> = 12.71*)					
Wife	420	65	13	14	8
Husband	219	71	7	14	8
Both	279	58	15	15	12
Total	918	64	12	14	9

<sup>a</sup>The frequencies of problems in all responses (never, seldom, sometimes, and often) were compared in the chi-square test.

\*Significant at the five percent level.

The problem of enough food was found significantly associated with income index, disposable income, net income, and perceived adequacy of income. Of these income measures, perceived adequacy had the highest association with the problem of "not enough food until there was money to buy more." Both perceived adequacy of income and enough food were subjective measures whereas the other measures were objective. By observing the perceived adequacy, the food problem could be better predicted by observing the other measures which were similar in their ability to explain variance in the food problems and to predict its occurrence.

Net income, the measure used currently for determining food stamp eligibility, was associated with the food problem but did no better in explaining variation or predicting the problem than did disposable income and income index which are easier measures to calculate than net income.

Families who never had the food problem were twice the poverty threshold (206 income index), had \$8543 mean disposable income and had \$7555 mean net income. These families on the average thought they could afford necessities only but not luxuries. Differences in means occurred for families who never had the problem and the other families. Fine distinctions of seldom, sometimes and often were not so helpful in establishing differences.

The findings offer some validation of the established poverty line. At income levels comparable to those at which public assistance was available, families in this study had problems of enough food until there was money to buy more. Almost half the poor families (income index under 100) and one third the near-poor families (income index of 100 through 124) had the food problem "often" or "sometimes". The average income index for those "never" having the problem was twice the poverty line (206) and for those "sometimes" having the problem it was at the near-poverty line (125). By the simple calculation of disposable income in \$1000 levels the percentage of those who "often" or "sometimes" had the food problem was 59, 46, 41, 36, 33, 23, and 19 beginning with the under \$1000 level to the \$7000 and above level. Exactly 33 percent of the families reporting "often" or "sometimes" were at the 100-124 income index level which, for a family of four, was comparable to the \$4000-4999 disposable income level where exactly 36 percent reported the problem this frequently.

Current benefits from food stamps are designed to enable a family to purchase the Economy Food Plan, although the estimated percentage getting an adequate diet on this Plan is very low. At the level of 100 income index which uses the cost of the Economy Plan in its calculation, 40 percent reported "sometimes" or "often" having enough food.

If families experienced the problem of not enough food they tended to experience other financial problems. Prediction of the food problem was increased about 50 percent by observing any one of the problems. About half the families who "often" had the problems of rent or house payment, utilities, or equipment maintenance, also "often" had the food problem. About half of the families who "often" had the food problem also "often" had problem of kids' things, doctor or medicine, clothing and personal care, and savings. Assuming money was allocated in order of the most pressing needs and, therefore, those needs were not allowed to become problems as much as other areas, the frequency of reported problems indicated their importance to these families. Having enough food was fourth following the more pressing needs of utilities, rent, and equipment maintenance but before the need to meet large bills, expenditures for doctor or

medicine, clothing, kids' things, and saving. The families that had a greater percentage of income committed to credit had the greater percentage reporting the food problem.

The problem of enough food for families under 125 income index, comparable to those eligible for food stamps, was clearly associated with their perceived adequacy of income, objective adequacy of income (income index), net income, and source of income but not with disposable income or income dependability. Additional associations were all the other financial problems, ethnic status, number of years family had been formed, husband's age, oldest child's age, husband's occupation, family type, socio-economic status, and person mainly responsible for using the money. Not associated with frequency of the food problem were percent of income committed to fixed expenses, language spoken, homemaker's age, homemaker's occupation, husband's occupation, homemaker's education, husband's education, who mainly decides the use of money, and frequency of the family's eating together. Many of these variables were associated with the food problem for other income index groups.

Families that were over-represented in "often" having the food problem tended to have income indexes of under 150, disposable incomes of under \$7000, perceived their income as not adequate or meeting necessities only, not dependable income, and income from only social security, welfare or welfare plus some earnings. They tended to be black or Spanish-American, one-parent families with a female head either employed or non-employed and in the lowest socio-economic class. In two-parent families, the husband tended to be under age 18 or over 55 and semi-skilled with both husband and wife handling the money.

Findings suggest that for lower-income families inadequate income, objective and perceived, caused the food problem but for higher income families such variables representing management and expectations as income dependability, percent of income committed to fixed expenses, language spoken, homemaker's and husband's occupations, women's responsibility as family head or wife, and homemaker's education was related to the food problem.

#### REFERENCES

- Bleckman, B.M. et al. Setting National Priorities: The 1975 Budget. (Washington D.C.: The Brookings Institute 1974.)
- Bickel, G.W. and MacDonald, M. Participation Rates in the Food Stamp Program: Estimated Levels, by State. (University of Wisconsin-Madison, January 1975)
- Participation Rates in the Food Stamp Program: Estimated Levels for 1974, by State. (Washington D.C.: Bureau of Social Science Research, Inc., Feb. 1975.)
- Hollingshead, A.B. Two Factor Index of Social Position. (1965 Yale Station, New Haven, Conn.: Author, 1957.) (Mimeographed)
- MacDonald, M. Food Stamp Program Regulations. (Institute for Research on Poverty, Madison, Wisconsin: University of Wisconsin, 1975) (Mimeographed).
- Manning, S.L. Financial Management Practices of Families with Steady or Fluctuating Incomes, Doctoral dissertation, Cornell University, Ann Arbor, Michigan: University Microfilms, 1960, No. 60-02086.
- Orshansky, M. "How Poverty is Measured." Monthly Labor Review, 1969, 92(2), 37-41.
- Patterns of Living Related to Income Poverty in Disadvantaged Families: A Basebook. (Ames: Iowa Agriculture and Home Economics Experiment Station Report No. 74, 1974.)
- Peterkin, B. and Clark, F. "Quality of Diets at Cost Levels of USDA Food Plans." Family Economic Review, 1969 (Sept.), 7-8.
- Projector, D. and Weiss, G. Survey of Financial Characteristics of Consumers. Federal Research Technical Paper. (Washington, D.C.: Board of Governors, U.S. Federal Reserve System, 1966.)
- U.S. Department of Commerce, Bureau of Census. Maps of Major Concentrations of Poverty in Standard Metropolitan Statistical Areas of 250,000 or More Populations, (Vol. 1, 3. Washington, D.C.: U.S. Government Printing Office, 1966.)

ATTITUDES TOWARD  
ENERGY CONSUMPTION: SEGMENTING  
THE GASOLINE MARKET

David J. Barnaby and Richard C. Reizenstein  
University of Tennessee

Abstract

This paper provides an empirical assessment of consumer attitudes and behavior relative to gasoline usage during the Arab oil embargo. The gasoline market is segmented into usage groups by AIO's and socioeconomic variables. Information included in segment descriptions suggests differential marketing strategies and tactics for each, directed toward reducing consumption and conserving energy.

Introduction

For most of the twentieth century, American consumers have been accustomed to abundant supplies of inexpensive energy. In addition, a myriad of products and services provided by seemingly limitless sources of power have fostered an "energy affluent" style of life. This almost unconscious dependency on an unending energy supply came to an abrupt and disconcerting halt with the Arab oil embargo in late 1973. Suddenly, unexpectedly, the prospect of energy allocation, rapidly rising prices, and shortages became a reality, with gasoline rationing a very real possibility -- all within an extremely short time span. The public began to realize that those who had been forecasting just such conditions, those who had been labeled "doomsayers", were, in fact, forecasting a situation which could fundamentally alter the American life style: the energy crisis.

The sudden energy shortfall created a need for information on public awareness of, attitudes toward, and usage of world energy resources. The limited economic data available did not provide insights into the opinions and consumption patterns of various segments of the American populace. Thus, a people-oriented information need exists, in order to provide inputs to government, business, and consumer groups, and to assist in policy formulation, product and promotion development, and conservation planning.

Overview

Unfortunately, this need for people-oriented data on the energy crisis specifically related to attitudes toward, perceptions of, and usage of energy resources, has primarily remained unfulfilled. Major public opinion organizations such as Opinion Research Corporation have conducted weekly polls, but these are primarily longitudinal analyses of a limited number of key categories; they do not deal with in-depth exploration of consumer attitudes and perceptions [Opinion Research Corporation, 1974]. Talarzyk and Omura, on the other hand, have performed in-depth analyses of consumer attitudes, perceptions, and usage patterns related to energy; however, their data base is a cross-sectional study conducted in March, 1974, after the lifting of the Arab oil embargo [Omura and Talarzyk, 1974; Talarzyk and Omura, 1974]. The current study attempts to establish a baseline of relevant attitudinal and usage information anchored at a key point in time (February, 1974, at the peak of the oil embargo). Data is thus provided for the cross-

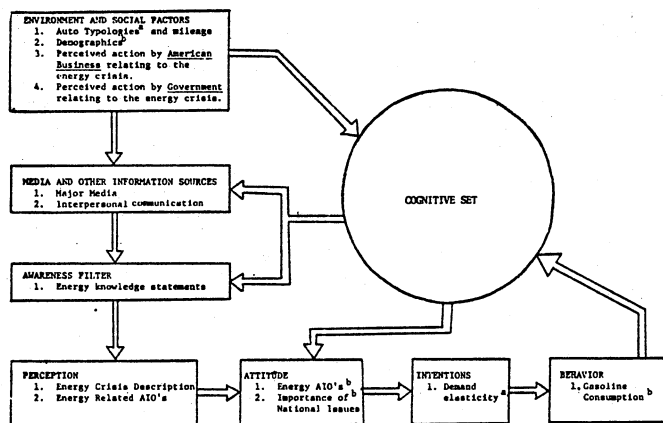
sectional analysis to be discussed below; in addition, it yields a basis for future longitudinal analyses.

Although the intent of this research is primarily to describe varying attitudes and energy consumption patterns among users in the Southeastern United States, a conceptual framework specifying various stages in the decision process has been included to show the range and relationship of the variables measured. A general paradigm, presented in Figure 1, incorporates the types of variables used to characterize group behavior and defines the relationships between sets of variables. This description of the consumer decision process borrows heavily from and paraphrases the Engel, Kollat and Blackwell model [Engel, Kollat and Blackwell, 1971].

Briefly, the general model of the gasoline decision process formulated in Figure 1 suggests that several levels of variables are important in shaping gasoline consumption behavior. These influences include the constraints of environmental and social factors as determined by sources of information. Informational influences in turn may be filtered according to environmental factors which will affect perceptions. Within the individual's cognitive framework, attitudes shaped by perceptions and environmental factors may lead to behavioral intentions and ultimately to overt behavior which could provide feedback through various levels of the decision process.

FIGURE 1

FRAMEWORK FOR GASOLINE DECISION PR. 2255



<sup>a</sup>Only in November, 1974 study, forthcoming.  
<sup>b</sup>These variables are the focal point of this study.

For the purpose of this research only, portions of three of these sets of variables were included for analytical purposes. Within the environmental and social factor variable set, 15 demographic variables were used. In the attitudinal variable set, 31 AIO variables and 10 importance rankings of national issues were included in the analyses. Finally, a surrogate for behavior was measured by reported gasoline

consumption. Given the interaction between variables shown in Figure 1, the framework for the analysis of the data was predicated on testing the relationship between behavior as measured by gasoline consumption, and demographic and energy/pollution related attitudinal items.

#### Methodology

In February, 1974, at the peak of the Arab oil embargo, a mail questionnaire was sent to 2500 residents of three medium-sized (100,000-350,000 population) Southeastern cities. As the questionnaire dealt with the major issues of air pollution and air pollution-energy tradeoffs as well as the energy crisis, test cities were selected on the basis of the only universal measure of air pollution, the average number of suspended particulates of matter per cubic meter of ambient air. Thus, one low pollution city (Columbus, Georgia), one medium pollution city (Charlotte, North Carolina), and one high pollution city (Chattanooga, Tennessee) were chosen.

The total of 2500 questionnaires was distributed proportionate to the population of each of the three cities, based on 1970 census figures. Respondents were selected within cities from the telephone directory, using a systematic random sampling technique. Specially designed envelopes, colorful commemorative stamps, and a standard follow-up letter were all employed to maximize returns [Reizenstein and Barnaby, 1975]. A total of 2389 questionnaires were ultimately delivered (111 were undeliverable due to moves or insufficient address), resulting in 922 usable responses, a rate of 38.6%.

#### Analysis and Results

##### Importance of National Issues

To determine the relative importance of a set of national problems and issues, ranking of the importance of these issues were used to establish a framework of "consumer concern" against which specific energy data could be compared. Table 1 lists the various issues by order of mean ranking for a sample of 922 subjects. While the three most important national issues include "the energy crisis", it appears that this issue is perceived as somewhat less important than either "corruption in government" or "inflation". Given the timing of the survey, February, 1974, at the apex of the Arab oil embargo, it is interesting to note that two other issues are perceived to be of greater importance by the sample respondents. This fact is somewhat in contradiction to Gallup data<sup>1</sup> which found the same three problems to be of greatest importance, but in inverse order; it may, however, be indicative of the lesser effect of the energy crisis on the Southeast as opposed to metropolitan areas across the nation.

##### Energy and Pollution AIO's

In addition to the global problem areas detailed in Table 1, a more specific set of 31 variables dealing directly with energy and pollution attitudes, interests and opinions was included in the analysis. The AIO's were scaled on a five-point interval from 1, strongly disagree, to 5, strongly agree. (AIO statements were phrased and ordered to attempt to minimize any halo effects.) Results of the responses to the 31 AIO's

are shown in Table 2. The energy related AIO's indicated various levels of disagreement with regard to the necessity of gasoline rationing, tax incentives for energy producers, and reduction of late evening radio and television programming. Various levels of agreement were shown for statements indicating the oil companies' deliberate creation of the energy crisis to enhance profitability, lessening of pollution from reduced automobile travel, negative effects on the economy from loss of jobs, and changes in our traditional life style.

Table 1  
IMPORTANCE RANKINGS  
(February, 1974)

Issues	Mean Rankings
Corruption in Government	2.38
Inflation in the U.S.	2.52
The Energy Crisis	2.98
Racial Tensions	3.68
Middle East Tensions	4.29
Drug Abuse Problems	4.30
Organized Crime	4.35
Environmental Pollution Problems	4.61
Unemployment	4.84
Poverty	5.11

Table 2  
ENERGY AND POLLUTION RELATED AIO'S  
MEANS AND STANDARD DEVIATION

Variable	Mean	Standard Deviation
1. Most people should voluntarily stop using products that pollute our surroundings.	4.0	1.3
2. Rationing of energy resources will be necessary for at least the next five years.	2.6	1.4
3. The reduction of air pollution is more important than energy conservation.	2.4	1.4
4. At this point in time our traditional energy resources (coal, oil, natural gas, etc.) are insufficient to continue energy consumption at the present rate.	3.1	1.6
5. Special consideration (tax reduction, etc.) should be given to those companies involved in the production of energy (utilities, petroleum companies, coal companies, etc.).	2.2	1.5
6. The energy crisis should be viewed as beneficial since it will reduce auto travel and the accident rate will decrease.	2.9	1.5
7. Because of the abundance of coal, industries should be encouraged to switch to coal as a fuel despite the air pollution it causes.	2.7	1.5
8. I find it difficult or unpleasant to breathe around here because of air pollution.	2.3	1.4
9. Air pollution is not as much of a problem in this city as in others I have visited.	3.8	1.5
10. The energy crisis has been created to distract the public from other pressing national issues.	2.9	1.6
11. Most people don't really care about pollution problems.	3.3	1.5
12. There is too much television time wasted on public service announcements relating to pollution control.	2.0	1.3
13. People do not cause pollution, industry does.	2.0	1.3
14. One benefit associated with the energy crisis will be a reduction in air pollution due to fewer automobiles on the road.	3.5	1.3
15. We should forget about reducing pollution until our energy problems are solved.	2.4	1.5
16. Our drinking water has a peculiar taste because of industrial wastes in our rivers and streams.	2.2	1.4
17. The oil companies did not deliberately create the energy crisis to increase their profitability.	2.3	1.4
18. The government should impose rigid controls to force companies to reduce their pollution.	3.7	1.4

<sup>1</sup>Gallup Poll, February 15-18, 1974 [Opinion Research Corporation, 1974].

Table 2  
(Page 2)

Variable	Mean	Standard Deviation
19. There is very little that an individual can do about pollution.	1.9	1.3
20. The energy shortage will hurt our economy due to the loss of jobs in energy related industries.	4.3	1.2
21. Because of the energy shortage, no television or radio programming should be permitted after 10:00 in the evening.	2.1	1.4
22. We should continue to support the Israeli position in the Middle East despite the energy problems caused by the shut-off of Arab oil.	3.2	1.6
23. Many fish and shell-fish can no longer be eaten because of industrial wastes in streams, rivers, lakes, and oceans.	4.0	1.2
24. An energy shortage will not affect our way of life.	1.4	1.0
25. Companies that advertise their efforts to reduce air pollution are more concerned with public relations than pollution control.	3.4	1.4
26. Gasoline rationing would be an unnecessary hardship on our citizens.	3.5	1.6
27. It is often quite hazy downtown due to air pollution.	3.3	1.5
28. Due to the recent introduction of anti-pollution devices on new cars, cars are no longer a serious source of air pollution.	2.3	1.3
29. The energy crisis is a pressing national problem which demands real sacrifices on the part of all citizens.	3.7	1.5
30. Why worry about cars causing pollution; it is the trucks and buses that cause the real problem.	2.2	1.3
31. It should be illegal to keep one's house warmer than 68° in winter.	1.8	1.2

In an effort to determine the extent of any overlap in the variety of AIO's and to provide a possible structure for the constructs being analyzed, a factor analysis was performed.<sup>2</sup> Varimax rotation of the principal components solution yielded eight factors with eigenvalues greater than one. The eight factors explained more than 53% of the variance of the original 31 variables. (Therefore, the number of variables were reduced by a factor of 4 with a loss of less than one-half of the original information.) The varimax rotation successfully spread the variable loadings across the eight factors so no cross-loadings of variables were evident, and enhanced the interpretation of the factor. An arbitrary cut-off of .4 for variable loadings was used in preparing Table 3.

In Table 3 descriptive names are attached to the eight factors consistent with the variables having the highest loadings. Factor 1, "Pollution is of Secondary Importance", implies, once again, that energy issues outweigh pollution problems, given the existing "personally negative effects of pollution" (Factor 2). A "gloomy, pessimistic" dimension regarding energy and pollution is shown by Factor 3 as opposed to a "silver lining" dimension in Factor 7. Factors 4 and 6 underscore the primary importance of pollution and energy, respectively, while Factor 8 indicates a necessity for gasoline rationing in the immediate future.

Although the interpretation of the factors suggests a general structure on which the variety of AIO's may be grouped, further analysis using the factor scores generated from this program will provide descriptors of various segments of the gasoline market. The inclusion

<sup>2</sup>The purpose of this factor analysis is considerably different from a similar analysis of energy related to AIO's by Talarzyk and Omura [Talarzyk and Omura, 1974]. Their purpose was to ascertain respondent logic and AIO item association.

of factor scores along with demographic variables to attempt to segment the gasoline market via discriminant analysis is discussed below.<sup>3</sup>

Table 3  
FACTOR ANALYSIS OF ENERGY AND POLLUTION AIO'S

Variable Loading	
	<b>1. Pollution is of Secondary Importance (15.7%)</b>
0.642	We should forget about reducing pollution until our energy problems are solved.
0.586	There is too much television time wasted on public service announcements relating to pollution control.
0.463	Because of the abundance of coal, industries should be encouraged to switch to coal as a fuel despite the air pollution it causes.
0.432	Most people don't really care about pollution problems.
	<b>2. Pollution Affects Me Negatively (9.6%)</b>
0.683	I find it difficult or unpleasant to breathe around here because of air pollution.
0.607	It is often quite hazy downtown due to air pollution.
-0.505	Air pollution is not as much of a problem in this city as in others I have visited.
	<b>3. I am Generally Pessimistic About the Effects Of Energy-Pollution Issues (7.4%)</b>
0.723	The energy shortage will hurt our economy due to the loss of jobs in energy related industries.
0.553	Many fish and shell-fish can no longer be eaten because of industrial wastes in streams, rivers, lakes, and oceans.
0.418	The energy crisis is a pressing national problem which demands real sacrifices on the part of all citizens.
	<b>4. Pollution is of Primary Importance (5.2%)</b>
0.586	The reduction of air pollution is more important than energy conservation.
0.449	The government should impose rigid controls to force companies to reduce their pollution.
0.436	Most people should voluntarily stop using products that pollute our surroundings.
	<b>5. Pollution is Not My Fault (4.4%)</b>
0.655	Why worry about cars causing pollution; it is the trucks and buses that cause the real problem.
0.524	Due to the recent introduction of anti-pollution devices on new cars, cars are no longer a serious source of air pollution.
	<b>6. Energy Problems are of Primary Importance (4.2%)</b>
0.645	The oil companies did not deliberately create the energy crisis to increase their profitability.
0.530	Special consideration (tax reduction, etc.) should be given to those companies involved in the production of energy (utilities, petroleum companies, coal companies, etc.).
0.435	At this point in time our traditional energy resources (coal, oil, natural gas, etc.) are insufficient to continue energy consumption at the present rate.
	<b>7. The Energy Crisis has a Silver Lining (3.7%)</b>
0.587	The energy crisis should be viewed as beneficial since it will reduce auto travel and the accident rate will decrease.
0.442	One benefit associated with the energy crisis will be a reduction in air pollution due to fewer automobiles on the road.
	<b>8. Gasoline Rationing will be Necessary</b>
0.528	Rationing of energy resources will be necessary for at least the next five years.

<sup>3</sup>For a related analysis including sources of information as predictor variables relating to energy usage, see Barnaby and Reizenstein [Barnaby and Reizenstein, 1975].

Table 3  
(Page 2)

Variable Loading	
-0.586	Gasoline rationing would be an unnecessary hardship on our citizens.
(Σ = 53.4%)	

<sup>a</sup>The SPSS Factor Analysis Program was used in performing the analysis (Sample size equals 922).

### Gasoline Usage Segment Description

The dependent variable under investigation in this study is reported gasoline consumption during a critical period of gasoline supply. Weekly gasoline usage data, collected on a series of "fuel gauge" scales, provided a measure of the distribution of gasoline consumption. In Table 4, frequency of gasoline usage per week is shown for a range of five gallon increments. More than two-thirds of the 922 respondents report using between 6 and 20 gallons of gasoline weekly, with a non-uniform distribution over this range. Approximately 18% of the sample used more than 25 gallons per week, while over 7% reported zero consumption.

Table 4  
GASOLINE USAGE FREQUENCY RESPONSES

Gallons/Week	Percentage	N
0	7.2%	66
1 to 5	4.4%	41
6 to 10	23.4%	215
11 to 15	22.2%	204
16 to 20	20.5%	188
21 to 25	4.4%	41
26 to 30	6.8%	63
31 to 35	2.2%	20
36 to 40	8.3%	77
41 or greater	.6%	7
	100.0%	922

For purposes of discriminant analysis, three consumption groups were arbitrarily formed to represent "light", "medium" and "heavy" gasoline consumption segments of the market. The three groups included consumption ranges of less than 10 gallons per week, 10-19 gallons per week and 20 or more gallons per week. As shown in Table 5, the groups are of unequal sizes with group 1 the smallest group, and group 2 the largest. Approximately a two-thirds sample of the original 922 respondents was selected for analysis. The remaining 307 members of the total sample were withheld from analysis and used as a hold-out sample to test the predictive ability of the discriminant functions. All groups showed better than 50% correct classification despite greatly unequal group sizes.

The BMD and Cooley and Lohnes [Cooley and Lohnes, 1971] multiple discriminant analysis programs were used to test the ability of a set of 24 variables (8 AIO variables computed from factor scores and 16 demographic variables) to discriminate among the three levels of gasoline consumption. The results of this analysis are presented in Table 5. Since the Cooley and Lohnes program is not a stepwise routine, the BMD program was used initially to determine the relative importance of those predictor variables that maximally separate usage groups. Of course, the extent of any correlation among the AIO variables is minimized because of the use of factor scores to represent orthogonal combinations of the original AIO variables. However, the demographic

variables include a considerable overlap of information. Therefore, those predictor variables entering the analysis first may mask the effects of other correlated variables and may impair statistical significance. Table 5 includes the four AIO's and the seven demographic variables that are found to be statistically significant at the .05 level in discriminating among the three gasoline usage groups. The 13 non-significant variables are included as a footnote to the table.

Table 5  
A DISCRIMINANT ANALYSIS OF THREE LEVELS OF GASOLINE CONSUMPTION

Variable <sup>a</sup>	F-Ratio <sup>b</sup>	Group 1	Group 2	Group 3
		Less Than 10 gal/wk.	10 to 19 gal/wk.	More Than 19 gal/wk.
		N	Mean (Std. Dev.)	Mean (Std. Dev.)
#12 Number in household	22.45	2.27 (1.68)	3.03 (1.25)	3.42 (1.58)
#23 Income	22.38	2.29 (1.80)	3.32 (1.63)	3.64 (1.66)
#3 I am generally Pessimistic and skeptical about energy-pollution issues	19.47	3.92 (1.76)	4.70 (0.88)	4.69 (1.00)
#19 Length of Residence	13.15	3.23 (2.01)	4.15 (1.41)	3.90 (1.47)
#1 Pollution is of secondary importance	10.17	2.35 (1.43)	2.90 (1.08)	2.95 (1.07)
#21 Age	9.10	3.71 (2.23)	4.52 (1.57)	4.23 (1.39)
#22 Education	8.12	3.01 (1.98)	3.67 (1.47)	3.72 (1.41)
#14 Family member 6-14 years	6.70	0.28 (0.73)	0.39 (0.73)	0.59 (0.89)
#7 The Energy Crisis has a Silver Lining	6.12	2.00 (1.21)	2.41 (1.02)	2.23 (0.95)
#8 Gasoline rationing may be necessary	4.77	-0.33 (0.99)	-0.56 (1.17)	-0.74 (1.17)
#17 Working Spouse	3.28	1.67 (1.47)	2.06 (1.34)	1.91 (1.21)

<sup>b</sup>Univariate F-Ratio Calculation based on 2 and 612 Degrees of Freedom;  $F_{.05, 2, 120} = 3.07$

Multivariate F-Ratio for 50 and 1176 Degrees of Freedom = 3.38;  $F_{.005, 60, 00} = 1.53$

<sup>a</sup>Variables not entered include factors 2, 4, 5, 6 and the following demographic variables: Sex, Marital Status, Family members 0-5 years, Family members 15-19 years, Family members more than 20 years, Occupation, Mobility, Race, and City.

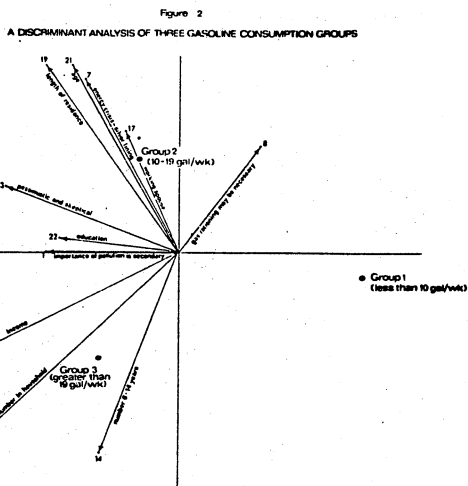
Those factors that individually provide statistical separation among the groups include: (1) "Pollution is of secondary importance"; (3) "I am generally pessimistic and skeptical about energy-pollution issues"; (7) "The energy crisis has a silver lining"; and (8) "Gasoline rationing may be necessary". In addition, (12) family size; (14) families with children 6-14; (17) families with working spouses; (19) length of residence; (21) age; (22) education; and (23) income are descriptors that provide group separation.

Moreover, Table 5 gives a relative indication of differences across groups for each variable by presenting significantly different group means. For instance, group 3 is composed of individuals with relatively higher incomes than group 2, which in turn includes members with average income levels greater than group 1. Conversely, group 1 sees a greater necessity for gasoline rationing than group 2 or group 3. While a variable by variable comparison of group differences is provided in Table 5, an alternative method of displaying this same data which enhances the total description of each group is presented in Figure 2.

The Cooley and Lohnes program calculates N-1 discriminant functions among N groups. Hence, in our three group study, two discriminant functions are calculated. By plotting the scaled eigenvectors, and the three group centroids of the two functions, a spatial representation of descriptors and group centroids is possible in two-space [Johnson, 1971 and Massey, 1965]. This configuration of descriptor variables as vectors and group centroids shows a "picture" of the gasoline



market and aids description of each group via the set of significant variables determined through the discriminant analysis. For example, in Figure 2, group 1, which is the low consumption group, is quite isolated from the other two groups; all descriptor vectors with the exception of the factor, "Gasoline rationing may be necessary", "load" less on this group than on groups 2 and 3. Group 2 gasoline consumers can be characterized by the proximate descriptor vectors as older, longer length of residence, greater likelihood of a working spouse, pessimistic and skeptical, and see the energy crisis as providing a silver lining. The high consumption group, 3, can be described as more educated with more children 6-14, a higher income level, and more working members of the family who generally see pollution as of secondary importance.



Such descriptions of consumer groups with various levels of gasoline consumption may provide insights into the appropriate marketing strategies required to reduce the energy consumption of the medium and heavy usage groups. Of particular importance is the selection of descriptor variables that provide not only a basis for segment identification, but suggest informational and attitudinal idiosyncracies of each segment to aid in the formulation of appropriate marketing programs in an attempt to alter traditional consumption patterns. A mechanism for differentially reaching various segments with unique characteristics may be derived from the inclusion of informational descriptors as suggested in [Barnaby and Reizenstein, 1975]. Obviously, any number of additional socio-economic variables, such as the gasoline consumption constraint inherent in ownership of a large domestic automobile, will provide useful information in discriminating among gasoline user groups.

#### Conclusions and Recommendations

The discriminant analysis of the gasoline consumption groups provides a vehicle for identifying distinct characteristics associated with such segments. These characteristics may be used as a basis for understanding group composition and the formulation of strategies directed toward reduction of excessive consumption where appropriate and feasible.

Group 1, the low and non-usage segment (less than 10 gallons per week), possesses the strongest positive attitude toward gasoline rationing. These individuals would appear to be younger, to have fewer children in the household, to have fewer working spouses, to have

lower family incomes, to be less educated, and to be shorter term residents of the community. Therefore, it would seem highly likely that these individuals are young singles or married couples in the early stages of the family life cycle. This group might include owners of smaller imported and subcompact automobiles as well as those who use mass transit and car pooling to meet their transportation needs. (This hypothesis will be tested in future analyses.) Since this group is presently a low consumption segment, a continuation strategy is deemed advisable for them. Advertising should emphasize the continued existence of energy problems, highlighting solutions in use by individuals whose characteristics match those of users and non-users within this group. A conservation orientation should be featured, targeted toward both owners and non-owners of automobiles. Ads directed toward the former might highlight gasoline saving tips, while those directed toward the latter could cite availability, convenience, and cost saving aspects of mass transit and car or van pooling.

Group 2, the medium usage segment (10 to 19 gallons per week), seems to have the attitude that the energy crisis has a silver lining. This attitude is reinforced by their generally skeptical outlook, another evidence of which is expressed in their opinion that pollution is of secondary importance. This group is older, more educated, long term community residents, and reasonably affluent (in part due to a working spouse). They seem to be skeptical about environmental issues which are heavily promoted in media as being of national importance; they require concrete, irrefutable evidence before their skepticism can be overcome. Therefore, a data-based communication strategy is recommended to promote decreased gasoline consumption within this group. Documented evidence from sources recognized by the public as having a high degree of credibility,<sup>4</sup> could be used as a focal point of advertising and publicity directed toward this segment. The composition of this segment must thus be analyzed prior to any such campaign to determine which information sources would be most credible to them.

Group 3, the heavy usage segment (greater than 19 gallons per week), is least inclined toward gasoline rationing and think that pollution is of secondary importance. They are characterized by a large number of people in the household emphasizing the 6 to 14 age range. They, therefore, fall into a "full nest" stage of the family life cycle, where the parents spend a great deal of time (and gasoline) acting as chauffeurs for the children. This group, being the highest income segment of the three (as well as the highest education segment), can afford to buy the additional gasoline required for such activities. Like group 2, this segment is pessimistic and skeptical about energy and pollution issues.

Due to the similarity of attitudes between 2 and 3, a data-based communication strategy is also deemed appropriate for the latter. The tactical thrust, however, should be oriented toward the future in light of the family structure of group 3, emphasizing the effects of possible gasoline shortages on the future welfare of their children. This approach must stress the credibility of data sources, a procedure similar to that suggested for group 2. Naturally, there may be differences in source credibility between segments. Other tactics to reduce the group 3 consumption level, might be promotion geared toward car pooling and more

<sup>4</sup>The New York Times, The Washington Post, recognized authorities and certain civic and public service organizations are examples of highly credible information sources.

efficient trip routing as a means of saving gasoline while not radically altering current life style.

Finally, it should be noted that segments 2 and 3 represent 73% of the respondents analyzed, thus indicating that higher gasoline consumption was the rule rather than the exception even at the peak of the Arab oil embargo. In terms of cost-benefit analysis, therefore, the data-based communication strategy and supporting tactics should receive paramount attention. It is this strategy which would appear to have the greatest potential for reducing consumption in medium and high usage segments of the gasoline market.

#### References

David J. Barnaby and Richard C. Reizenstein, "Profiling the Energy Consumer: A Discriminant Analysis Approach," a paper presented at the National ORSA/TIMS Conference (Chicago: May, 1975).

William W. Cooley and Paul R. Lohnes, Multivariate Data Analysis. New York: John Wiley & Sons, 1971.

James F. Engel, David R. Kollat, and Roger D. Blackwell, Consumer Behavior, Second Edition. New York: Holt, Rinehart, & Winston, 1973.

Richard M. Johnson, "Market Segmentation: A Strategic Management Tool," Journal of Marketing Research, 8 (February, 1971), 13-19.

William F. Massey, "Discriminant Analysis of Audience Characteristics," Journal of Advertising Research, 5 (February, 1965), 39-40.

Glenn S. Omura and W. Wayne Talarzyk, "Relationships Between Consumers, Shopping and Leisure Activities and their Attitudes Toward the Energy Crisis: A Cross Sectional Study," a paper presented at the Association for Consumer Research Annual Conference (Chicago: October, 1974).

Opinion Research Corporation, Trends in Energy Consumption and Attitudes Toward the Energy Shortage, Volume V, Princeton, N.J.: Opinion Research Corp., 1974.

Richard C. Reizenstein and David J. Barnaby, "Assessing Methods of Increasing Response Rates in Mail Questionnaires," a paper presented at the National ORSA/TIMS Conference (Chicago: May, 1975).

W. Wayne Talarzyk and Glenn S. Omura, "Consumer Attitudes Toward and Receptions of the Energy Crisis," a paper presented at the American Marketing Association Fall Conference (Portland: August, 1974).

## CONSUMER SATISFACTION AS THE ULTIMATE LIFE FORCE

Clark Leavitt, The Ohio State University

### Abstract

The paper deals with the nature of satisfaction and its relation to pleasure in an information processing framework. Methodological problems in the measurement of Quality of Life are considered as well as the problems arising in use of subjective social indicators for policy decisions.

Satisfaction as a phenomenon is important for several reasons. It is--or at least should be--a key concept in any theory of behavior. As I have suggested elsewhere (Leavitt, 1975) it represents an opportunity to make the field of consumer behavior more than just another area of applied psychology. Perhaps most important of all: satisfaction is the criterion of the quality of life. This paper explores the implications of the last statement.

Most scholars would agree that we all strive for some kind of positive state. Many terms have been used to label it: gratification, pleasure, positive hedonic tone, happiness, joy, enlightenment, self-actualization, nirvana and many others. The term satisfaction has the advantage of being widely used and philosophically neutral. Four underlying questions concerning satisfaction will be discussed in order to quell the rise of unreasonable expectations. I note at the outset that the most important question--How do we achieve this most desired state?--will not be discussed. The four questions are:

1. Is satisfaction pluralistic in some sense? If so, how does consumer satisfaction fit with other sources?
2. Is satisfaction measurable?
3. What are the methodological problems and caveats to be considered in measuring satisfaction?
4. What should we do with the results?

This is the logical order of these questions. However, I would like to begin by discussing the last question--What do we do with our measures of satisfaction after we have invested the time and effort to work them out? I see two broad possibilities: general monitoring of the state of the organization, either the entire social system or some part of it such as the corporation, and program evaluation: that is, assessment of planned experimental intervention. The first use is very much tied up with the social indicators movement and I would like to first review this movement and place consumer satisfaction in that framework.

### Social Indicators and Quality of Life

Historically, the development of social statistics in the United States relied upon individual investigators developing *ad hoc* estimates in pursuit of their specific interests (Darmstadter, 1966). The first major departure from this tradition occurred in the thirties as a result of the government's mandate to adopt positive programs to facilitate economic recovery. To provide the information required for planning and assessment, the national income and product accounts, the nation's

first social statistics were devised (Kuznets, 1956), based primarily on the macroeconomic theories of Keynes and his students (Lewis, 1952).

The accounts were continually revised, elaborated and disaggregated in response to the changing nature of the major economic problems facing the nation: war mobilization, postwar recovery, recession and inflation (Lewis, 1952). While criticisms were leveled at individual measures on statistical or technical grounds (Kuznets, 1948), their applicability as a measure of the state of the nation, or the general welfare was seldom challenged until recently (Galbraith, 1973).

But a growing concern for other aspects of life has given rise to a doubt regarding the suitability of economic measures as sole indicators of the state of society.<sup>1</sup> The heightened salience of social, cultural and environmental goals, and their institutionalization in "Great Society" programs, brought the need for data portraying the nation's status and rate of progress along these dimensions into bold relief (Gross, 1966).

In response a voluminous literature has developed in a very brief period (Wilcox, 1972). A major portion of this body of work is devoted to conceptual and theoretical considerations (Bauer, 1970). The definitions, attributes and hazards of a system of social measurements have been widely debated. Investigators have suggested the use of objective, and more recently, subjective indicators to measure and monitor the quality of life. Legislation has been proposed which would give such assessments equal footing with the national income and product accounts.

Objective Indicators. Early attempts to use objective indicators were primarily of an illustrative nature. Scholara proposing that a certain aspect of life be measured demonstrated the feasibility of their positions by discussing examples of available data that assessed, or could serve as a surrogate for, the concept under consideration (e.g., Thorndik, McHale, 1967). More recently, the objective approach has been extended to more fully recognize the need for multiple dimensions. Cross-sectional studies have compared the quality of life in different countries (Rummel, 1969), states or regions (Liu, 1973) and most frequently, urban areas (Flax, 1972). The status of various segments of the population has been studied, including racial and ethnic minorities (Rokeach, 1970), women (Ferriss), youth, the aged (Todd, 1970) and others.

Terleckyj, in a most comprehensive work, sets forth nineteen national goals related to six major areas of

<sup>1</sup>U.S. President, Research Commission on National Goals. Goals for Americans: Programs for Action in the Sixties. Englewood Cliffs, N.J.: Prentice-Hall 1960. Also, Bauer, Raymond A. "Data Needs of a Science for Solving Social Problems," Address to the 1968-69 Seminar Series, The Travelers Research Center, Inc. November 7, 1963. These developments find precedent in U.S. President, Research Commission on Social Trends. Recent Social Trends in the United States. New York: McGraw-Hill, 1933. However, the recommendations of this commission, which was chaired by William Ogburn, were largely lost amongst the pressing concerns of economic issues.

concern: Health and Safety; Education, Skills and Living Standards; Income; Economic Equality; Human Habitat; and Art, Science and Free Time. After describing existing statistics measuring the current status in each area, activities and programs facilitating progress toward each objective are discussed. Estimates are then made for the next decade of feasible accomplishments and the costs of these increments.

While advancing our knowledge of the process of social measurement, the objective approaches are subject to several major criticisms. First, there is a tendency for the measures selected to reflect a programmatic, or institutional bias. The dimensions actually considered are most often suggested by existing data which, in turn, reflect the goals and priorities of established social, bureaucratic and political elites. This inhibits judgment as to the completeness or correctness of the goal set itself. Further, while purporting to measure a wide range of life's aspects, many quantitative systems of indicators, implicitly, consider only the distribution of economic resources (James, 1972).

A second criticism, closely related to the first, is that goals and measures are most frequently examined in isolation. There is often no scheme for weighing general goals or the individual statistics taken as indicative of progress towards these objectives (Mitchell). Exceptions are found in Terleckyj's attention to overall budgetary constraints, in the factor analytic studies by Gitter and various collaborators, of crime (Gitter, 1971), and education (Gitter, 1970) and in Wilson's discussion of socio-economic indicators.

Finally, the very essence of quality of life measurement would seem to be the determination of the goals, aspirations, hopes, fears and satisfactions of the major segments of the population (Maslow, 1972). Once identified, the researcher's task becomes the measurement of the level and rate of progress of society along these dimensions. While objective data can provide certain insights, only perceptual and behavioral measures are capable of fully defining the relevant variables and generating the desired assessments.

#### Social Experimentation

The use of quality of life indicators in a system of "social accounts" clearly stems from an economic tradition embodying a very different attitude toward data than that of psychologists. On the other hand, there is a long tradition of program evaluation in which more concern is shown for the kinds of measurement problems that experimental psychologists are more likely to take seriously.

But the divergent values and emphasis stem ultimately from the question of which is the cart and which is the horse. The economist seems to operate on the assumption that a system manager, if he has enough information about the state of the system, can anticipate and avoid problems.

The opposite point of view has been eloquently expressed by Campbell beginning with his paper, "Reforms as Experiments" (1968). He has taken the position that society needs to develop the habit of regarding its problem solving efforts as tentative and in need of experimental evaluation. This suggests an entirely different use of satisfaction measures and an entirely different set of methodological concerns. Primarily, I have in mind those discussed by Campbell and Stanley (1965) under the rubric of quasi-experimental design.

#### Subjective Indicators and Consumer Studies

Subjective measures of the quality of life have received far less attention in the literature than their objective counterparts. However, many studies of attitudes and beliefs related to individual problem areas can advance our understanding of the quality of life concept. Recent examples include studies of attitudes towards the use of violence as a change mechanism (Blumenthal, 1971), job satisfaction (Herrick, 1972), the utility of housing (Shinn, 1972), citizen and consumer satisfaction (Lingoes, 1972), and various environmental issues. Returning to our original list of four questions, I would like to consider to what extent satisfaction is plural and many-faceted and where consumer satisfaction fits in. Works such as those just mentioned add to the substantive knowledge concerning individual aspects of the quality of life and, frequently, make major methodological contributions. However, they do not attend to the problem of weighting in aggregating individual facets according to their importance.

This weighting task has been approached in a limited number of more broadly focused studies. Bettman's investigations of citizen's preferences for social consumption (Bettman, 1971) and Sawchuk and Gitter's study of eight dimensions of the quality of life (Sawchuk, 1971) might be mentioned. But like those employing objective indicators, they are characterized by ex ante specification of the variables to be considered. Discovering the overall dimensionality of the quality of life concept and making an empirically-based assessment of the representativeness of the set of measures employed are strategies largely ignored. There are, however, three notable exceptions:

1. Post-treatment interviews with the students responding to the instrument used by Sawchuk and Gitter indicated that the HEW aspects examined in the study formed an incomplete description of the quality of life (Gitter, 1971). A content analysis was performed and the sixteen aspects of life shown in Table 1 resulted. This set of dimensions has been studied in two subsequent works.

TABLE 1 (From Gitter, 1971)

#### Quality of Life

- 
1. Availability and utilization of leisure time and recreation
  2. Educational attainment
  3. Enjoyment of art in all forms (music, sculpture, dance, etc.)
  4. Finding fulfillment and satisfaction with work and life
  5. Getting along with people in general
  6. Getting along with people who are close to you
  7. Health
  8. Income
  9. Living and working physical environment
  10. Morality: Values and conduct in dealing with others
  11. Opportunity and ability to "get ahead" or advance oneself through education and work

TABLE 1 (Continued)

12. Opportunities for involvement in community, state and national affairs and ability to influence these affairs
13. Personal physical appearance
14. Rights: free speech, assembly, privacy, equality before the law, etc.
15. Safety of self and property from crime
16. Safety of self and property from international conflict

Gitter and Franklin collected ratings for a variety of target persons using these dimensions. The scores were averaged, across targets, and factor analyzed. Four

factors explaining 65% of the variance resulted: Physical Context, Security, Aesthetic-Intellectual, and Compatibility.

In a parallel effort Gitter and Knoche obtained ratings of the importance of the same set of dimensions (Gitter, 1971). Again the results were factor analyzed. Three factors were derived explaining the 65% of variance in the original data: Social Acceptability, Security, and Aesthetic-Intellectual.

Unfortunately, Gitter and Knoche do not present the results of a four factor rotation of their data. However, the general comparability of the two sets of results can be seen from Table 2. Item 10, Morality and Values, was not seen by the authors as showing significant loadings in either study. Each of the other dimensions appears in one of the analyses with a significant loading. The greatest difference between the two seems to be the emergence of the "Physical Context" factor in the Gitter and Franklin Study.

TABLE 2

Comparison of the Gitter and Franklin  
and The Gitter and Knoche Item Loadings  
(Loadings greater than 3 retained)

	Study						
	Gitter and Franklin (quality)				Gitter and Knoche (importance)		
	Security	Intell- Aesthetic	Compat- ibility	Phys. Context	Security	Aesthetic- Intell.	Social Accept.
1. Leisure		-.80				.74	
2. Education		-.82				.73	
3. Art		-.79				.76	
4. Satisfaction with work		-.49				.74	
5. People in general			-.81				.66
6. People who are close			-.88				.78
7. Health			-.38	.67	-.47	-.38	.46
8. Income				.77	-.43		.66
9. Phys. Env.				.78	-.53		.56
10. Morality & Values			.61	.43		-.35	.64
11. "Get Ahead"		-.51		.47			.74
12. Involvement				.68	-.50		.44
13. Pers. App.				.49			.75
14. Rights	.61			.47	.84		
15. Safety from Crime	.85				.84		
16. Safety from Water	.82				.77		
Variance Explained	13%	17%	14%	21%	21%	19%	25%

TABLE 3

## Importance of Dalkey and Rourke Dimensions

1. Love, caring, affection, communication, interpersonal understanding; friendship, companionship; honesty, sincerity, truthfulness; tolerance, acceptance of others; faith, religious awareness.	15.0
2. Self-respect, self-acceptance, self-satisfaction; self-confidence, egoism; security; stability, familiarity, sense of permanence; self-knowledge, self-awareness, growth.	11.5
3. Peace of mind, emotional stability, lack of conflict; fear, anxiety; suffering, pain; humiliation, belittlement; escape, fantasy.	10.0
4. Sex, sexual satisfaction, sexual pleasure.	9.5
5. Challenge, stimulation; competition, competitiveness; ambition; opportunity, social mobility, luck; educational, intellectual stimulating.	8.0
6. Social acceptance, popularity; needed, feeling of being wanted; loneliness, impersonality; flattering, positive feedback, reinforcement.	8.0
7. Achievement, accomplishment, job satisfaction; success; failure, defeat, losing; money, acquisitiveness, material greed; status, reputation, recognition, prestige.	7.0
8. Individuality; conformity; spontaneity, impulsive, uninhibited; freedom.	6.0
9. Involvement, participation; concern, altruism, consideration.	6.0
10. Comfort, economic well-being, relaxation, leisure; good health.	6.0
11. Novelty, change, newness, variety, surprise; boredom; humorous, amusing, witty.	5.0
12. Dominance, superiority; dependence, impotence, helplessness; aggression, violence, hostility; power, control, independence.	3.5
13. Privacy	2.0

2. Dalkey and Rourke employed a panel of students in an iterative procedure designed to discover and rank the relevant dimensions. Initially, a large number of open-ended responses were obtained from the subjects. These were grouped by the authors into 48 categories. Similarity ratings between pairs of categories were obtained and the best of dimensions was reduced to 13 factors through a cluster analysis of these ratings. Finally, a Delphi procedure was employed to generate importance ratings for these factors. The resulting factors and their importance ratings are shown in Table 3.

3. By far the most ambitious project reported is that of Andrews and Withey (1974). They developed the components of life satisfaction on a much more adequate basis than the two previous studies. They then used the components to predict an overall satisfaction rating. This procedure positions the consumer satisfaction component in an adequate, and representative framework, thus offering the best estimate of the importance of consumer related satisfactions in the narrow sense of commercially available products and services. Most areas of satisfaction are consumption related however, (see Table 4).

## Methodological Notes

By answering the first question we have certainly answered the second: Is satisfaction measurable? It is measurable in many ways. What happens when they dis-

agree? Is some overall scale the final arbiter?

Ultimately, the only answer here can be an experimental one. If people find themselves in an obviously better life which of their subjective responses is most sensitive to this fact? This suggests a program of validation in which an array of measures are used under field conditions to study this impact of actual change done perhaps on an experimental scale.

## Applications

What do we do with the results? I have already pointed to two types of use for the social system at large: trend indicators for general monitoring and indicators program evaluation for use in experimentation.

Perhaps more important is our need to understand how satisfaction actually works as a feedback servo-mechanism to control productive and marketing efforts. Both Stokes (1975) and Rosenberg (Czepiel, Rosenberg and Akerele, 1974) have discussed the broad policy implications as well as the methodological entailments. Sturdivant and Ginter (1975) at Ohio State have discussed the problem from the point of view of the corporation. We (Leavitt, 1974) have formulated a micro-analytic behavioral model of the marketing communication system which postulates a specific relation between satisfaction and repurchase behavior. It is in expanding and testing this kind of formulation that a sensitive measure of satisfaction may find the greatest relevance.

TABLE 4<sup>a</sup>  
 Perceived quality of life (Life #3) predicted  
 by subsets of the 30 domains

N = 1297

Data sources: May 1972 national surveys

Domain subset:		Top 16	Top 6	Selected 12
Percent variance explained:		54%	49%	52%
Multiple correlation:		.73	.70	.72
Population estimate:		51%	48%	50%
		<u>beta</u>	<u>beta</u>	<u>beta</u>
14+20+22	EFFICACY INDX	.27	.28	.25
12	YOURSELF			
1+2+3	FAMILY INDX	.18	.17	.19
4	FAMILY			
83+85	MONEY INDX	.15	.20	.15
27	AMOUNT OF FUN	.16	.21	.16
87	HOUSE/APARTMENT	.12	.13	.11
6	THINGS DO W FAMILY	.09	.10	.08
38	TIME TO DO THINGS	.09		.07
123	YOUNG PEOPLE THINK	.08		
30	SPARE TIME ACTIVITIES	.08		.08
112+115	RECREATION INDX	.06		
105+107+108+109	NATL GOVT INDX	.08		.09
106+107+109	NATL GOVT INDX			
97+99+101	*CONSUMER INDX	.06		.06
101	GOODS & SERVICES			
93+103	LOCAL GOVT INDX	.06		
74	HOUSEWORK	.07		
116+117	MEDIA INDX	.05		
7	YOUR HEALTH	.06		.06
75+76+77+79+80	JOB INDX			.02
75	YOUR JOB			

<sup>a</sup>From Andrews and Withey (1974)

\* The three consumer items are:

1. The way you can get around to work, schools, shopping etc.
2. The doctors, clinics and hospitals you would use in this area.
3. The goods and services you can get when you buy in this area - things like food appliances, clothes.

To reverse the old saw: "If something can be measured it exists." This may reflect best the present state of the art--a state of neglect.

#### References

1. F. Andrews and S. Withey, "Developing Measures of Perceived Life Quality," Social Indicators Research, v. 1, 1974.
2. Raymond A. Bauer (Ed.), Social Indicators (Cambridge, Mass., MIT Press, 1966).
3. James R. Bettman, "Measuring Individuals' Priorities for National Goals: A Methodology and Empirical Example," Policy Sciences II (1971), pp. 373-390.
4. James R. Bettman and Robert B. Andrews, "Social Marketing and Consumer's Preference for Social Consumption," Proceedings, Association for Consumer Research (1972), pp. 679-688.
5. Monica D. Blumenthal, et al., "Attitudes Towards Violence," Ann Arbor, Michigan: The University of Michigan Institute for Social Research, 1971.
6. Angus Campbell and Philip Converse, The Human Meaning of Social Change, (New York: Russel Sage Foundation, 1972).
7. Donald T. Campbell and Julian C. Stanley, "Experimental and Quasi-Experimental Designs for Research in Teaching," in N. L. Gage (Ed.) Handbook of Research on Teaching (American Educational Research Association, 1963).
8. Donald T. Campbell, "Reforms as Experiments," American Psychologist (1967).
9. Donald T. Campbell, "Artifact and Control," Reprinted from R. Rosenthal and R. L. Rosnow (Eds.), Artifact in Behavioral Research, Academic Press, (1969), pp. 351-352.
10. Donald T. Campbell, "Or the Genetics of Altruism and the Counter-Hedonic Components in Human Culture," Journal of Social Issues, Vol. 29, No. 3, 1973.
11. Donald T. Campbell and Philip Brickman, "Hedonic Relativism and Planning the Good Society," Ab-

- stract of a lecture for the Adaptation - Level Conference, University of Massachusetts.
12. D. T. Campbell and H. L. Ross, "The Connecticut Crackdown in Speeding," Time series data in quasi-experimental analysis, in E. R. Tuftee (Ed.), The Quantitative Analysis of Social Problems.
  13. J. A. Czepiel, L. J. Rosenberg and A. Akerele, "Perspectives on Consumer Satisfaction," paper presented at American Marketing Association Educators' Conference, Portland Oregon (1974).
  14. Joel Darmstadter and Sidney Sonenblum, "The National Accounts and Emerging Social and Economic Problems," Proceedings, Business and Economic Statistics Section, American Statistical Association.
  15. Amitai Etzioni and Edward W. Lehman, "Some Dangers in 'Valid' Social Measurement," The Annals of the American Academy of Political and Social Science, CCCLXXXII (September, 1967), pp. 1-15.
  16. Abbott L. Ferriss, Indicators of Trends in the Status of American Women (New York: Russell Sage Foundation, 1971).
  17. Michael J. Flax, A Study in Comparative Urban Indicators: Conditions in 18 Large Metropolitan Areas (Washington, D.C.: The Urban Institute, 1972).
  18. John Kenneth Galbraith. The New Industrial State, Rev. Ed., 1973.
  19. A. George Gitter and S. Franklin, "Subjective Quality of Life Indicators of Sixteen Aspects of Life," Communication Research Center Report No. 58, 1971.
  20. A. George Gitter and E. Knoche, "Importance Ratings of Sixteen Aspects of Life," Communication Research Center Report No. 59 (Boston: Boston University, 1971).
  21. A. George Gitter and R. R. Peterson, "Toward a Social Indicator of Education - A Pilot Study," Communication Research Center Report No. 44 (Boston: Boston University).
  22. Robert A. Gordon, Business Fluctuations, (New York: Harper, 1952, pp. 28-57).
  23. Bertram M. Gross, "Planning: Let's Not Leave it to the Economists," Challenge, (September-October, 1965), pp. 30-33.
  24. Bertram M. Gross, Social Intelligence for America's Future: Explorations in Societal Problems, (Boston, Allyn and Bacon, 1969).
  25. Neal Q. Herrick and Robert P. Quinn, "The Working Conditions Survey as a Source of Social Indicators," Monthly Labor Review (April, 1971), pp. 15-24.
  26. Jack M. Hicks and Donald T. Campbell, "Zero-Point Scaling as Affected by Social Object, Scaling Method, and Context," Journal of Personality and Social Psychology, V. 2, No. 6, pp. 793-808. (1965)
  27. Richard D. James, "Measuring the Quality of Life," The Wall Street Journal, (May, 1972).
  28. Simon Kuznets, "National Income: A Summary of Findings," (New York: National Bureau of Economic Research, 1946).
  29. Simon Kuznets, "Discussion of the New Department of Commerce Income Series," Review of Economics and Statistics, XXX (August, 1948).
  30. Clark Leavitt, "Research Project Proposal: Construction of a Quality of Life Scale," The Ohio State University, College of Administrative Science, (1973).
  31. Clark Leavitt, "Strong vs. Weak Effects of Mass Communication: An Empirical Test," in G. D. Hughes and M. E. Ray (Eds.), Buyer/Consumer Information Processing, Durham University of North Carolina Press, 1975.
  32. John P. Lewis and Robert C. Turner, Business Conditions Analysis, 2nd ed., (New York: McGraw-Hill Book Company, 1967), pp. 13-53.
  33. James C. Lingoies and Martin Pfaff, "Measurement of Subjective Welfare and Satisfaction," a paper presented at the Annual Meeting of the American Economic Association jointly with the Association for the Study of the Grants Economy, December 1971, New Orleans, Louisiana.
  34. Ben-Chieh Liu, et al., The Quality of Life in the United States 1970: Index Rating, and Statistics, Kansas City, Mo.: Midwest Research Institute, 1973.
  35. Abraham Maslow, "The Good Life of the Self Actualizing Person," Moral Problems in Contemporary Society: Essays in Humanistic Ethics, Paul Kurtz Ed., (Englewood Cliffs, N.J.: Prentice-Hall, 1969).
  36. John McHale, "Science, Technology and Change," The Annals of the American Academy of Political and Social Science 371 (May, 1967, pp. 120-139).
  37. Mancur Olson, Jr., "Social Indicators and Social Accounts," Socio-Economic Planning Science II, (1969), pp. 334-346.
  38. Anita B. Pfaff, "An Index of Consumer Satisfaction," Proceedings, Association for Consumer Research (1972), pp. 713-737.
  39. Milton Rokeach and Seymour Parker, "Values as Social Indicators of Poverty and Race Relations in America," The Annals of the American Academy of Political and Social Sciences (March, 1970), Vol. 388, pp. 97-111.
  40. Richard Ruggles and Nancy D. Ruggles, National Income Accounts and Income Analysis 2nd ed., (New York: McGraw-Hill Book Company, 1956).
  41. Rudolph J. Rummel, "Indicators of Cross-National and International Patterns," The American Political Science Review (March, 1969), pp. 127-147.
  42. R. Sawchuk and A. George Gitter, "Eight Subjective Indicators of Quality of Life," Communication Research Center Report No. 51, (Boston: Boston Univ., 1971).
  43. Elanor B. Sheldon and Wilbert E. Moore, Indicators of Social Change: Concepts and Measurements, (New York: Russell Sage Foundation, 1968).
  44. Allen M. Shinn, Jr. "Measuring the Utility of Housing: Demonstrating a Methodological Approach," Social Science Quarterly (June, 1972), pp. 88-102.
  45. Raymond C. Stokes, "Consumerism and the Measure-



- ment of Consumer Dissatisfactions," in Attitude Research Bridges the Atlantic, Philip Levine, (Ed.), American Marketing Association (1975).
46. Frederick D. Sturdivant and James L. Ginter, "Assessing Social Performance: A Preliminary Model," Division of Research Working Paper Series, College of Administrative Science, The Ohio State University, (1974).
  47. Nestor E. Terleckyj, National Goals Accounting (Draft #1), Washington D.C.: National Planning Association, 1972, pp. 17-63.
  48. Sheldon Todd and Jacqueline Anderson, "A System of Social Indicators for the Aged." A paper prepared for the Administration on Aging by the Operations Research and Policy Systems Division, American Rehabilitation Foundation, Minneapolis (1970).
  49. John P. Wannous and Edward E. Lawler, III, "Measurement and Meaning of Job Satisfaction," Journal of Applied Psychology, 56 (1972), pp. 95-105.
  50. Leslie D. Wilcox, et al., Social Indicators and Societal Monitoring: An Annotated Bibliography, (San Francisco: Jossey-Bass Inc., 1972).
  51. John Oliver Wilson, "Quality of Life in the United States--An Excursion into the New Frontier of Socio-Economic Indicators," symposium sponsored by the Environmental Protection Agency, pp. 260-288 (not dated).

## CS/D: THE PROGRAM PLANNING AND EVALUATION PERSPECTIVE

H. Keith Hunt, Brigham Young University

### Abstract

The application of consumer satisfaction/dissatisfaction measures to program planning and evaluation is discussed. Early usefulness is predicted in creating recognition of consumer dissatisfactions. Comparing the worth of programs or evaluating programs' impacts seems a bit in the future.

### Introduction

Since 1972, there has been some interest among a few ACR members in the development and use of consumer satisfaction and dissatisfaction measures. Public policy is one of the more frequently mentioned application areas for these measures. This paper discusses the potential uses and limitations of consumer satisfaction and dissatisfaction measures in the program planning and program evaluation (PP&E) dimensions of public policy decision making.

### Program Planning and Evaluation

Let me lay the groundwork for this discussion by explaining briefly the rationale underlying the program planning and evaluation (PP&E) approach to public policy decision making. The current emphasis on PP&E supposedly started in the Department of Defense with Robert McNamara as the originator. The Department of Defense and many other federal branches and agencies have the mind-boggling problem of funding many more projects than a cadre of analysts and monitor and evaluate on a continuing basis. The sheer magnitude of the comparisons coupled with the highly specialized expertise needed to evaluate many of the projects, demanded that the projects be under decentralized management. Yet the top decision maker or decision team was responsible to congress and/or the president and through them to the public, to be sure all the projects devouring public funds were worth the costs they were incurring. The Department of Defense tackled this problem and developed the management framework we know today as PP&E. The PP&E approach centers on the idea that the key elements in any organization are its programs, where a program is a unit of work which in itself is a whole unit and which differs from other units of work being directed and performed in the same office or division. The work of any division can be summarized by listing its current programs. And, it naturally follows, the work of any division can be evaluated by evaluating the progress made in its set of programs during the relevant time period. Budgets are developed and explicated in a program framework, allowing budget review groups to apply a cost-benefit analysis in assessing whether, given the proposed budget and expected outcomes, a program was beneficial or not.

The key activities in program planning are the stating of the objective or objectives, the stating of the activities that would be undertaken to accomplish those objectives, and the stating of how the impact of those activities would be evaluated in assessing program effectiveness.

For each program a document is prepared which explains the basic problem or quest which is the reason for the program's existence. Based on this statement of the problem, the objective or set of objectives is stated, followed by the set of activities to be done during specified time periods at specified costs to accomplish the objectives. And, finally, a research procedure is specified which will investigate the degree to which the activities accomplished the objectives.

The use of programs as the basic unit of structure in the resource allocation process and the requirement that each program's impact be evaluated have brought an increased interest in methods and techniques for explaining the basic importance of programs and for evaluating the programs' impacts. Many methods have been proposed. One particular frame of analysis is that known as CS/D, consumer satisfaction and dissatisfaction.

### Relevance of CS/D for PP&E

CS/D has potential relevance for program planning and evaluation in three different ways: (1) in forcing initial recognition of the problem, (2) in evaluating the seriousness of the problem, and (3) in evaluating the impact of the program activities on accomplishing the program's stated objectives.

### CS/D: Forcing Initial Recognition of the Problem

Throughout this discussion I will be using the letters CS/D to stand for a methodology which will be general enough to (1) measure consumer satisfaction and dissatisfaction in product categories at a specific enough level to be useful to public policy groups and business management groups for evaluating areas needing change and (2) broad enough for use as a social indicator of consumers' senses of well being and adequateness in the national consumption system. While such a methodology does not now exist in operational form, I think progress is being made at a sufficient rate to expect that within the next two or three years such a methodology will be on-line.

Most agencies and departments become aware of problems within the organization's domain of responsibility from two basic sources: (1) complaint data and (2) internal research and monitoring of trends. Those who use only the complaint data are essentially reactive organizations; those using both the complaint data and internal problem identification efforts are proactive organizations. CS/D is useful in both these areas.

The monitoring of complaint data is still disorganized in most complaint areas, often lacking even a single agency to attend to it. Still, letters from irate citizens, from congressmen (usually initiated by irate constituents), and from organizations' own staff members' experiences, form the primary sources of complaints. Recent years have seen increased interest in whether this informal and biased complaint source could

possibly be supplanted by an ongoing complaint gathering mechanism. Discussion of this topic usually leads to the suggestion that what we really need is longitudinal research into consumer dissatisfaction--research that would not only peg dissatisfaction at a certain level in the current period but would also show trends in dissatisfaction within and across product categories. The catch to this line of thinking is that we haven't had a methodology or a conceptual framework for gathering and analyzing such data. Ralph Day and Laird Landon have done substantial work on this problem. They are reporting on their work in the next session.

One of the great advantages of CS/D over the current complaint process is that CS/D would enable us to estimate the actual number of dissatisfied persons, and even their differing levels of dissatisfaction, where the current process merely has a series of letters reaching agencies, departments, and congressmen demanding change, with no estimate of overall dissatisfaction being possible from them.

Another advantage of using CS/D is that it gets away from the bias inherent in the current complaint system where a very small number of complainers heavily influence the reactive efforts of federal organizations in ways that might not be in the interests or desires of most of the uninvolved noncomplainers. The problem, however, is whether a survey of CS/D would do a better job than the current haphazard process or whether it would merely replace one faulty system with another.

Beyond the complaint data issue, the second source of problem recognition is the internal research and monitoring efforts of some federal units. Rather than waiting for massive complaints to react to, some federal units are searching out problem areas while those areas are still developing. CS/D in its longitudinal use would certainly provide rich information for this early identify-and-correct effort.

Increased consumer satisfaction can come by increasing the actual benefit to the consumer or by taking actions to decrease negative benefits. Business usually performs the function of creating better products and thus creating increased satisfyingness in a product category. Government has the policing task of reducing unsatisfyingness due to such problems as safety, health, fraud, and malfunction. So while I keep talking about consumer satisfaction and dissatisfaction, we need to keep in mind that it seldom is government's function to increase the satisfyingness of an already satisfactory product. And if business had been willing to decrease the dissatisfaction of its products there would be no need for government to get involved in the matter at all. So in talking about program planning and evaluation issues, the federal system is primarily a dissatisfaction reducer, not a satisfaction increaser.

#### CS/D: Evaluating the Comparative Seriousness of Problems

Given that CS/D might be a useful first indicator or alert that a product area may not be yielding as much satisfaction as consumers think it should, that alert must then be responded to with a thorough analysis to discern the real state of affairs. Is the product faulty? Is consumer perception or expectation faulty? Has advertising led consumers to expect more than any such product could possibly deliver? These and similar questions would be seeking insight into whether the problem area really is a problem worth public resources to solve, and indeed, whether it is even solvable. If

the problem is considered worth continued attention, then the question arises: of the problems we know of, which ones should we allocate our scarce resources to solving.

Is CS/D of potential use in evaluating problem/program areas to aid in determining to what extent each should be funded? At present I see little hope for any early success in this area. While we seem to be close to having an operational CS/D methodology, that methodology in its early form will be primarily designed to indicate problem areas, not for determining the relative worth of programs. It is interesting to note that at least one methodologist, John Eighmey of FTC, has, in pondering the problem, thought that perhaps joint measurement might allow such comparisons. But the operationalization of these comments is at least months away. So, while there is some hope that CS/D might soon be operational in the complaint gathering function, it isn't likely that this development or other developments will allow us to assess the relative worth of programs in the near future.

Paying particular attention to the Federal Trade Commission's possible application of CS/D, one of the definite changes we have seen in the FTC's operation in the last couple of years is the attention paid to cost-benefit framework. How does CS/D fit into the cost-benefit frame of analysis? CS/D as it is currently conceptualized does not directly provide cost or benefit information. If that continues to be the case then CS/D will be of little use to decision makers whose evaluations are cast in the cost-benefit framework. About the best that could be expected would be thinking by Staff and Commissioners about what the cost is to consumers of a less than adequately satisfactory product, and what the cost would be to all parties concerned to make the product as satisfying as consumers think it should be. It appears that in a cost-benefit analytical framework, CS/D would have little value except as an initial indicator of a problem area worth further investigation.

#### CS/D: Impact Evaluation

Perhaps some time in the future we will see an evaluation research component of a program plan which states that the objective of increased satisfaction (or decreased dissatisfaction) will be judged to be successful if the CS/D measurement in that category moves from the negative side of the scale to the positive. At the present time all we can say is that consumer complaints should be reduced substantially. Again, this dimension of CS/D in the program planning and evaluation area lacks development enough to be operational.

#### Conclusion

CS/D, a methodology for measuring consumer satisfaction and dissatisfaction, is developing. PP&E, a planning and evaluation tool for managing resource allocation to programs, is a common approach in managing complex and diverse activities, especially in the public policy area. CS/D seems most likely to be used in PP&E as the initiator of awareness of the existence and magnitude of consumer dissatisfaction. CS/D also holds promise for comparing the worth of alternative programs and evaluating programs' impacts, but the realization of this potential requires substantial development in the methodology.

## CONSUMER SATISFACTION: A NEGLECTED LINK?

Donald J. Hempel, Session Chairperson, University of Connecticut  
Larry J. Rosenberg, Discussant, New York University

### INTRODUCTION

Is consumer satisfaction a neglected link in the study of consumer behavior? An equivocal answer is found in the literature of the professional groups and academic disciplines represented at this conference. One can find many instances where the concept of consumer satisfaction is invoked as a basic consideration in the design and evaluation of marketing programs (Pratt, 1972; Kotler, 1975). The concept is widely accepted as fundamental to the definition of marketing and it is represented in most comprehensive models of buyer behavior (McCarthy, 1975; Sheth, 1974). However, there are significant limitations in the application of the concept that stem from problems in its definition and measurement (Bell and Emory, 1971). The absence of a generally acceptable operational definition of consumer satisfaction appears to result in neglected implementation of this critical concept. It is almost an axiom of administrative systems that those areas of concern which are not measured tend to become neglected.

### MEASUREMENT LIMITATIONS

Given the established importance of the consumer satisfaction concept, the underdeveloped status of its measurement is difficult to explain. The operational development of this concept is particularly relevant in a period of growing concern about consumer attitudes toward business and increasing rates of consumer complaints. In the absence of its precise definition and measurement, marketing systems must often be judged with highly subjective criteria, and both the acceptance and implementation of the marketing concept is limited by subjective appraisal.

The hesitant response of many firms to consumerism and social responsibility may occur because the two fundamental elements of the marketing concept--consumer satisfaction and profit--are both plagued by measurement problems. From the viewpoint of the manager, the greater measurability of profit performance and its prominence on the scorecard of business are likely to shift attention toward this concept with some consequent neglect of those aspects of consumer satisfaction which are not directly reflected in profit. Those who hope to implement the marketing concept are faced with the significant challenge of applying measures of consumer satisfaction which can be coupled with measures of profit to achieve a balanced evaluation of total performance. There is need for those who specialize in consumer studies to develop operational measures of consumer satisfaction which are analogous to the concepts of "good health" in the medical profession and "efficiency" in engineering. Perhaps such efforts will be considered the hallmarks of our development as a profession.

### A REVIVAL?

In recent years there has been a resurgence of interest in the measurement of consumer satisfaction (Lingoes and Pfaff, 1972; Pfaff, 1972; Czepiel, Rosenberg, and Akerele, 1975). These renewed efforts

appear to be stimulated by opportunities to use measures of consumer satisfaction as twin indicators of operational effectiveness and social performance. Further development of these measures as components of a quality of life index may provide an operational linkage between specific organizational units and the social system in which they function.

The papers presented at this session show that the link is not as neglected as first impressions indicate. They cast light on various facets of the market system, including the consumer, the corporation, the government, and the consumer advocate. We should first ask the question - What is the link that we are dealing with? Is it the link between the reality of consumer satisfaction and theory, that between theory and measurement in terms of consumer satisfaction or dissatisfaction and the quality of life, or the link between measurements and policy applications by government or business? Our speakers take a look at several of these links and it is the chain formed by all of them together that is going to give us a command of this topic.

### THE ULTIMATE LIFE FORCE

Clark Leavitt looks at consumer satisfaction as the ultimate life force. He asks some excellent questions about the nature of satisfaction, its measurability, problems with methodology and what we can do with the results. This is an excellent agenda and his views contribute towards answering these questions. Several concerns ought to be taken into consideration. The first is the danger of over-promise: should we look at consumer satisfaction as perhaps less than the ultimate life force, but one of several contributing factors to the life force? What is the meaning of consumer satisfaction in the context of Clark's view on the quality of life? Is it a separate sub-category - consumer concerns for goods and services - or does it parallel and purvey all of the other categories that are listed? Our feeling supports the later interpretation and suggests that we better watch our methodology to allow for this interdependency. The studies that are cited in terms of the quality of life seem to give only minor attention to consumer satisfaction. This may be an area where we as consumer researchers have to conduct our own agenda so that we can fully focus on its total contribution. Our egos may be jolted a bit when we see on the studies that have been cited how small a piece of the total quality of life action we seem to have.

### PPE PERSPECTIVES

Keith Hunt's paper examines the program planning and evaluation perspective of consumer satisfaction/dissatisfaction. If, as Keith puts it, there is increasingly a marriage between CS/D and PPE, the coverage in this paper seems to tell us a lot about the groom PPE; but the bride CS/D seems to be heavily veiled and off to the side. Since the usefulness of CS/D in this perspective is being questioned, we should have more knowledge about the state of the art in terms of measuring and applying CS/D. Who is working on it and

what approaches are they using? Many methods have been proposed to evaluate program impact of which CS/D is one. But what are the other methods, how do these several methods compare, and what is the complementary nature of various methods? This paper links consumer satisfaction and dissatisfaction together with the convenient shorthand CS/D, but how are they related? In terms of methodology and scale construction, are they at separate ends of the same continuum or are they on different continua such as very satisfied to not satisfied and very dissatisfied to not dissatisfied? This is a complex process and the two can exist at the same time. For example, one might be satisfied with the styling and the speed of an automobile, but dissatisfied with its safety and miles per gallon.

Another basic issue is what should we expect of our CS/D methodology? Can it satisfy the dual requirements that both Clark Leavitt and Keith Hunt have set forth--that it will simultaneously apply to specific products as well as serve as a social indicator? It seems to be a tall order to serve both as a general indicator that takes on a monitoring quality like the Gallup Poll to provide data from which we can make inferences, and as a specific product evaluator where we would expect more diagnostic capacity. What is the linkage between complaints and CS/D monitoring? Again, are they viable as separate measurements or do they comprise a total system in which the strengths and weaknesses of each can somehow be complemented? This leads to the issue of whether the government should react or anticipate. A reactive posture often takes on the view of a policeman who responds only after there have been victims. It does not minimize the actual suffering of consumer that results from poor choices or bad experiences with products and services, or prevent the further erosion of credibility in our economic system and in many of our social institutions. There is need to truly anticipate these shortcomings before they become serious problems; to make the government a guardian of the process rather than just a policeman.

#### POLICY PERSPECTIVES

Ray Stokes' paper provides some policy perspectives concerning consumer satisfaction from the vantage point of one who has experience with this issue in both business and government. His focus on the food industry is also very relevant--it allows us to evaluate an important system in terms of its business and government interaction. One issue raised is that many of the consumerism programs have not been very successful. Are these shortcomings the responsibility of the program planners who pressured for it and conceptualize it--perhaps the consumer advocates? Who implemented it in a minimal commitment way, without sufficient educational resources to bring the programs to the consumer in a forceful way? What is the appropriate time frame? The outcome might well be different over the long run with sufficient education. And who does the education? Is it just business or does government in its guardian role have a responsibility to contribute? Another issue is the information gap that is pointed out in the food companies not knowing the percentage of dissatisfied consumers who actually complain. Why haven't they found out? Why haven't they taken a look at what might be called the iceberg hypothesis--that the complainers are a distinct minority who have special skills in a division of labor sense. They may be the articulate few who reflect the views of the larger silent majority below the water line. In this analogy the consumer advocates might be the seagulls who perch themselves on the top of the iceberg and not only speak for the vocal complainers,

their immediate constituency, but also represent the drift of the entire iceberg. A third issue is the timidity of manufacturers to print on packages or to publicize in advertisements their money back guarantees. Are they afraid of too much response or does their silence contribute to the consumer not knowing his rights? Business can recognize these rights but if it's done in a quiet way it is equivalent to non-recognition.

Ray Stokes' paper praises competition in the food industry. It is interesting to probe why the consumer advocates have such little regard for this accomplishment of what appears to be an extremely efficient system of food distribution. Perhaps competition at these levels could turn out to be bad social responsibility. In other words, social responsibility may get lost between the cracks of vigorous competition. For instance, the product differentiation and the multitude of package sizes and descriptions are regarded as good by the standards of competition. In terms of social responsibility, have we truly estimated the confusion and therefore the anxiety and further erosion of credibility that has resulted on the part of the consumer?

#### CONCLUSION

These papers are significant forward steps in the construction of an agenda for further exploration of this important link in consumer research and in our marketplace. There are many more issues that can be raised. Among the topics in need of further research are the aggregation problems resulting from intra-family differences in CS/D, the generalization of CS/D with specific products to other parts of the marketing system and temporal variations in reported levels of individual satisfaction.

#### REFERENCES

- Bell, Martin L. and William, Emory C. "The Faltering Marketing Concept," Journal of Marketing, 35 (October 1971), 37-42.
- Czepiel, John A, Rosenberg, Larry J. and Alkerle Adebayo, "Perspectives on Consumer Satisfaction", in R.D. Curhan, editor, Marketing's Contribution to the Firm and to the Society, Chicago: American Marketing Association, 1975), 119-123.
- Kotler, Philip, Marketing for Nonprofit Organizations (Englewood Cliffs, N.J.: Prentice-Hall, 1975).
- Lingoes, James C. and Pfaff, Martin, "The Index of Consumer Satisfaction Methodology." Proceedings of the Third Annual Conference of the Association for Consumer Research (College Park, Md.: Association for Consumer Research, 1972), 689-172.
- McCarthy, Jerome E. Basic Marketing: A Managerial Approach (Homewood, Illinois: Richard D. Irwin, 1975).
- Pfaff, Anita B. "An Index of Consumer Satisfaction," Proceedings of the Third Annual Conference of the Association for Consumer Research (College Park, Md.: Association for Consumer Research, 1972) 713-37.
- Pratt, Robert W. Jr. "The Index of Consumer Satisfaction and Corporate Marketing Policy," Proceedings of the Third Annual Conference of the Association for Consumer Research (College Park, Md.: Association for Consumer Research, 1972) 742-5.
- Sheth, Jagdish N. Models of Buyer Behavior: Conceptual, Quantitative & Empirical (New York: Harper & Row, 1974).

## COLLECTING COMPREHENSIVE CONSUMER COMPLAINT DATA BY SURVEY RESEARCH

Ralph L. Day, Indiana University  
E. Laird Landon, Jr., University of Colorado

### Abstract

Although widely relied upon by policy makers, the consumer complaint data used in the past have been seriously flawed. This paper reports a pilot study utilizing survey research methods for providing higher quality data on complaints and the behavior of complainers plus more general data on the entire range of the satisfaction dissatisfaction continuum.

### Introduction

Both consumer protection agencies and business firms still receive most of their feedback from consumers in the form of volunteered complaints from individuals who have experienced extreme levels of dissatisfaction with products or services. This paper will be concerned primarily with two issues with respect to complaint data: (1) What is the best way to collect it?; and (2) Shouldn't we be collecting and using something else instead, i.e., data on the entire range of satisfaction and dissatisfaction? These two questions will first be discussed in general terms, then a particular study utilizing a questionnaire initially designed for the Federal Trade Commission will be briefly described, and the paper will conclude with some suggestions for future research.

### The Quality of Volunteered Complaints

It is widely recognized that data on specific complaints reported by individuals and collected and summarized by various agencies are seriously flawed. Problems with complaint data are usually attributed to the fact that the people who come forward with complaints are not a representative sample of the population. In the words of Dr. Raymond C. Stokes, former Director of the Consumer Research Institute, "No one maintains that complaint letters which flow to governmental agencies or to private business firms come from a representative sample of the consuming public. In fact, the contrary is known to be true. Analysis of complaint letters show that they are heavily weighted by two groups: (1) people with time on their hands; and (2) highly educated, articulate people" (Stokes, 1974). Although the limited number of studies of consumer complaining behavior do not all agree on the precise ways in which complainers are different, there seems to be general agreement that people who come forward with complaints are quite atypical of the population (Warland, 1974).

In studies where researchers have gone to the consumer to get information on consumer complaints through survey research, one can break out those who have not complained because they have not been dissatisfied from those who have not complained even though they have been highly dissatisfied. The results of one recent study offered additional evidence that complainers are quite different and provides some new insights on the transmission of complaints. The study was a telephone survey of a national probability sample of 1,215 consumers. All respondents were asked "lately, have you gotten good and mad about the way you were treated as a consumer?" A large majority of the respondents (64.5%) said they had not been upset about their treatment as

consumers, 24.2% said that they had been upset and had taken some action, and 11.3% reported that they had been upset but had taken no action. The three groups were quite different on demographic variables and in their attitudes toward government and business. The upset-complainers were at one extreme (young, affluent, activist liberals), the upset-noncomplainers were at the other extreme (older, poorer, politically alienated), while the non upset-noncomplainers fell in between on most measures. While hardly definitive, these results suggest that the young and highly educated are overrepresented in complaints although they may be quite atypical of the population in terms of their consumption patterns (Warland, 1974).

Perhaps the most interesting result of this study was the small fraction of those who reported complaining who had taken action that would be likely to contribute to the complaint statistics of consumer protection agencies. Of the 294 respondents who reported taking action, there were only 15 cases of reports to "Better Business Bureau, government official, TV station or newspaper" and 10 cases in which the action was "contacted a lawyer, filed a suit, went to small claims court." The great majority of the complainers had complained to the firm or firms involved, boycotted the seller, refused to pay, or simply complained to relatives and friends (Warland, 1974). Without a lot more information than is now available about the processes which determine which complaints eventually reach consumer protection agencies, the assumption that these complaints even roughly reflect the nature and scope of consumer concerns seems highly questionable.

It should be recognized that the relevant issue is whether reported and recorded complaints are a representative sample of "the population of legitimate complaints" and not whether the complainers are a representative sample of the general population in terms of demographic characteristics. In other words, there are factors other than what the individual is like which bear on the accuracy of conventional complaint data, and these are not necessarily predicted by demographic data. While a comprehensive "theory of consumer complaining behavior" remains to be formulated and tested, a preliminary effort will be made here to identify the major factors. At a rather general level, we will consider complaining behavior as a function of four factors: (1) the individual's propensity to complain when dissatisfied; (2) the individual's opportunities to become dissatisfied with products or services; (3) the opportunities available to the individual to obtain redress and/or register complaints; and (4) disparity in consumer knowledge. Each of these will be discussed very briefly.

### Propensity to Complain

Different people will react differently to identical experiences with goods and services. Depending on the circumstances of purchase and use, expectations, tastes, and personality factors; the reactions of different individuals to exactly the same purchase situation may vary from a high degree of satisfaction to extreme dissatisfaction. Likewise, different individuals experiencing a similar degree of dissatisfaction will vary with respect to their complaining behavior and with respect to the kind of action taken if complaining

behavior is triggered. The concept of "propensity to complain" seems to be a useful approach to describing different patterns of complaining behavior.

There is at least a suggestion in the data of the studies mentioned above that the propensity to complain is related to education, age, affluence, and attitudes toward government and business. However, it is also very likely that some of these variables are associated with each of the other three factors.

#### Opportunities to Become Dissatisfied

The product which has been the undisputed occupant of the number one position on complaint lists of consumer protection agencies for years is the automobile. Ownership and use of automobiles is almost universal in our society, the family car is essential to the day-to-day operations of most households, the cost of owning and operating the automobile(s) is a major item in the household budget, and the car is a complex of electrical and mechanical systems which the typical owner understands very poorly. In view of these factors, it is hardly surprising that complaints involving automobiles exceed those for any other product. The volume of transactions involving automobiles is so great and are generally so large in value that the automobile would probably retain its number one spot in terms of the sheer quantity of complaints even if the automobile industry and automobile repair industry were able to make dramatic improvements in the quality of the product and in repair services.

By contrast, many products are so low in value that dissatisfactions are overlooked and other products are important to the user but are used by only a small segment of the population. In such cases, it is possible that the rate of dissatisfaction with the product might be rather high among those who use it but the volume of reported complaints might appear negligible. Therefore, in evaluating the meaning of consumer complaint data one needs to have a measure of the product's rate of use over the population and some idea of other factors such as the product's economic importance, its social importance, its complexity, and its degree of "foolproofness."

#### Opportunities to Obtain Redress or Complain

Another factor which influences the quantity of complaints is the ease with which an individual can obtain redress locally and conveniently in the event of extreme dissatisfaction. In general, complaints are less frequent when the consumer deals with a local business. Then negotiations to obtain redress can be entered easily and a complaint results only if the seller should refuse to provide redress acceptable to the consumer. On the other hand, if the source of the good or service is physically distant or otherwise difficult to contact, then the complaint may be initiated through the local Better Business Bureau or other agencies and eventually show up in consumer complaint data.

This suggests that the channels of distribution used for a particular product could effect the extent to which it shows up in complaint data. Mail order and door-to-door channels would seem vulnerable to this effect. The difficulty in seeking redress by people who live in rural areas or in small towns also might tend to raise the incidence of complaints through public channels, although this could be offset by the absence of convenient agencies for the transmittal of complaints.

#### Individual Knowledge

Another factor contributing to the biases that have

been noted in complaint data is the wide disparity in consumer knowledge about products and services across the total population. The less knowledgeable consumer will be less able to judge product performance and evaluate the goods and services he consumes. Also, he will be unfamiliar with procedures for seeking redress and in registering complaints. The disparity of consumer knowledge is to a considerable extent a function of formal education but is also influenced by age, geographical factors, and ethnic factors. As a result, the poorly educated, the elderly, ethnic minorities, and the residents of more remote areas may tend to be underrepresented in volunteered complaint data.

While these four factors have been discussed as if they were independent of each other, it is very likely that there are interactions among them and perhaps there are additional factors that have been overlooked. In any event, it seems clear that there are many opportunities for biases and inaccuracies to enter traditional complaint data.

#### Compliant Data by Survey Research

An attractive alternative to volunteered complaint data is the periodic collection of consumer data using a national probability sample of consumers interviewed in their homes. By giving a representative sample of consumers an opportunity to report on their recent experiences, a far more accurate picture of the incidence and nature of highly unsatisfactory experiences can be obtained. This should provide a representative sample from the "universe of potential complaints" without all of the distortions which would be introduced by waiting for a self-selected subset of all dissatisfied consumers to go to the trouble and effort involved in submitting formal complaints. In particular, it would provide useful data on those who were dissatisfied but took no action, those who sought and obtained redress within the business system, and those who had taken some action but gave up without either obtaining redress or filing a formal complaint. Also, the direct contact with the respondents would facilitate getting more complete information of the particular circumstances of each unsatisfactory experience so that more complete as well as more accurate data are obtained.

Although survey research removes many of the problems of volunteered complaint data, it introduces some of its own. If the survey questions are framed in a very general way, then problems of incomplete or selective recall might be troublesome. If aided recall is utilized by providing checklists of products and services, the interview may become unduly long and result in refusals and nonresponse bias. Biases may also be introduced by the content of the list itself. Whether recall is aided or unaided, there may be problems with the subject's ability to relate recalled experiences to a specific time frame so that projections can be made to a specific period of time (e.g., number of complaints in 1975). The timing of a one-shot survey could also be troublesome because of the seasonal nature of the sales of many products and services. There are probably additional problems on top of the standard problems of sample and instrument design, not to mention the very substantial costs involved in doing large sample national surveys.

A small number of survey studies relating to consumer dissatisfaction and complaints have been reported in the literature. The Warland, et. al., study mentioned above was a general study using unaided recall and was made by telephone to a national sample of 1,215 consumers. The respondent was asked if she had become angry over a recent consumer experience and, if she said yes, what she had done about it. A significant new finding

of this study was that a substantial segment of the population reported being upset over consumer experiences but did nothing at all about it. Obviously, this segment of the population is overlooked if one depends on volunteered complaint data.

Other recently published reports of survey projects have been limited to some restricted class of products. For example, one recent study looked at the consumer's "need for redress" with respect to women's personal care products. A telephone study of 466 women in 19 cities obtained information on "the one problem remembered most clearly," the degree of annoyance, actions to obtain redress, and subsequent repurchase behavior. On aided recall, only 19% reported either returning the item to the store or contacting a manufacturer. They were much more likely to stop buying the brand and to complain to their friends (34%). Only about 1% reported complaining to the Better Business Bureau (Diener, 1975). A recent study of selected consumer durables in Columbia, South Carolina, used aided recall in personal interviews to obtain complaint data from a convenience sample of 1,024 households. An interesting aspect of this study was that an effort was made to document the process of complaining and obtaining redress. Of 243 who made an effort to obtain redress, 168 were classified as successful and 75 as unsuccessful (Thomas, 1975).

#### The "National Consumer Survey" Project

The remainder of this paper will be devoted to a discussion of the design and pilot testing of a survey research instrument to be used in a large national study of consumer satisfactions, dissatisfactions, and complaining behavior. The original questionnaire was designed by Ralph Day while he served as Consultant to the Office of Policy Planning and Evaluation of the Federal Trade Commission and was modified by both authors for the pilot tests made in Bloomington and Boulder during the fall of 1974. It is being revised for use with a large national sample in early 1976.

The project represents the first effort made by federal consumer protection agencies to seek better information for policy making purposes than that provided by volunteered consumer complaint data. The initial objective of the study was to obtain complaint data which would be more current and complete. The data collection instrument was designed with this objective primarily in mind, but it also provided for obtaining data on the degree of satisfaction as well as dissatisfaction. Before the pilot studies were made, sections were added to the questionnaires including data on the complaining behavior of consumers.

When the data collection instrument was being conceptualized, the FTC's Bureau of Consumer Protection suggested that for their purposes the major objective of the study should be the provision of a products-by-complaints matrix, summarizing the incidence of complaints of various kinds for each of a comprehensive set of product categories. This was adopted as the overriding objective of the first National Consumer Survey and, of course, had a major impact on the design of the survey instrument. As an aid to understanding, interpreting, and validating the results, it was decided to include attitudinal questions relating to various marketing practices, marketing institutions, government control of business, and the usual demographics. Also, a section was included on the priorities consumers feel should be assigned to various consumer protection and consumer education activities.

In order to structure a products-by-complaints matrix, it is necessary to establish a set of categories for

the products and a list of specific complaints or complaint categories. The classification system for complaint data used by the Office of Consumer Affairs, the FTC, and other agencies for their complaint data in the past did not seem satisfactory. The major problem was the mixing together of items that are purchased under such widely varying circumstances that the differences seem to outweigh any common thread in the classification. An example of this is that the classification "automobiles" in past complaint data tabulations includes not only new and used cars but such things as repair services, accessories, gasoline and oil, and automobile insurance. This mixes together both durable and nondurable tangible products, services, and intangibles which are acquired and used in quite different ways and under very different market circumstances. While no comprehensive system of classification will be without flaws, it seemed more logical for the purposes of the survey to organize the classification around the usual marketing classification system which begins with a breakdown of all products and services into durable products, nondurable products, services, and intangible products. For convenience, the service and intangible categories were combined to yield three major classifications. Within these three major categories from 6 to 9 "functional breakdowns" such as food, clothing, housing, and transportation were established, and within each of these subcategories 6 to 15 groupings of products which have similar patterns of acquisition and use were established. The original questionnaire contained a total of 198 of these specific categories and provided an opportunity to write in any item which the respondent might feel was omitted at the end of each of the 25 second-level categories.

It is hardly burdensome to ask the respondent to indicate a level of satisfaction or dissatisfaction by making a check on a three-point scale for each of the 223 possible categories. However, the instrument devised to learn the specific nature of any complaints and to obtain information on complaint action took as many as 47 items in the Boulder instrument. Thus, if these questions were supplied for each product category, this section of the questionnaire alone would require more than 10,000 questions. Obviously this would be impracticable and some means of simplifying the task was required. Rather than abandoning the disaggregated product categories and settling for higher levels of aggregations of dissimilar items, it was decided to keep the specific categories in the 200 range but to ask the respondent to complete the complaint questions for only the one item in each of the three major sections with which he or she was most dissatisfied. The task for the respondent is first to choose one of the following answers for each of the approximately 70 items in each of the three major sections.

- I rarely or never use items in this category.
- I am satisfied with items in this category.
- I am somewhat dissatisfied with items in this category.
- I am very dissatisfied with items in this category.

After completing each of the three major sections, the respondent is asked to review the ratings in the sections, the respondent is asked to review the ratings in the section and: (1) pick the three items which are the most satisfactory, (2) the three items which are the least satisfactory, and (3) the one item of the latter three which is the most unsatisfactory. Then, for this particular category, any complaints the respondent has about that product are checked off from a list of 15 to 18 specific complaints.

Thus the products by complaints matrix can be completed with each respondent being represented by the one item in each of the three product sections which was desig-



nated "the most unsatisfactory." The answers on the "complaint action" questions associated with each of the three product categories also provides the basis of a "products by complaint actions" matrix. In addition to these "focused" responses related to complaints, the initial screening questions provide an indication of the general degree of satisfaction or dissatisfaction for every respondent for every one of the product categories with which he has had a recent purchase experience. Although the final evaluation of the instrument as the basis of a rather monumental national effort to measure consumer satisfactions and dissatisfactions is far from complete, efforts so far suggest that it works at least reasonably well. Even with the adoption of the "funneling technique" for obtaining data for a products-by-complaints matrix, the questionnaire is still quite long. The typewritten draft runs 29 pages which might be reduced to perhaps 18 to 20 pages in printed form. However, our preliminary testing indicated that when the questionnaire is self-administered without time pressure, most people will complete it although a few will quit in midstream and others will complain a bit after completing it.

#### Some Results from the Boulder Study

During the fall of 1974, both authors conducted consumer surveys in their respective home cities, Boulder, Colorado, and Bloomington, Indiana. Both used versions of the instrument developed at the FTC, with some additional questions added. The Boulder study was larger, obtaining 275 usable responses while the Bloomington study obtained 150. Although analysis has not been fully completed for either study, it is further along with the Boulder data and the results given here will be from the Boulder survey. Both studies used multi-stage probability samples of households and preliminary checks of the demographic variables suggest that they are representative. While the two cities are not representative of the national population, they do not appear to be so atypical to be unsuitable for evaluating methods of research on consumer behavior. In both studies, the questionnaires were delivered by university students who answered respondents' questions and made appointments to return within a few days to pick up the questionnaire.

While both the Boulder and Bloomington studies had the development of a better survey instrument for a subsequent national study as an important goal, each was intended to be a useful study in its own right. Together they provide the first data base containing measures of consumer satisfaction and dissatisfaction over a comprehensive set of product categories and would appear to be the most exhaustive study of the complaining behavior of consumers yet undertaken. These data can be analyzed and interpreted without waiting for the completion of the national study still in preparation, although any national projections will have to await the full national study. Hopefully, the present results, once the analysis has been completed, can provide momentum in a gradually increasing effort to better understand some aspects of consumer behavior which have been under-researched in the past.

Any comprehensive presentation of the data from the Boulder study is not feasible here. The results of the respondents' three levels of evaluation of each of the approximately 200 product and service categories are quite voluminous and do not lend themselves to summarization in any compact way. Some information on the particular product categories which appeared most frequently on the "three most satisfactory" and "three least satisfactory" lists were reported in a previous paper (Day and Landon, 1975) but a much more extensive presentation of data than space allows would be re-

quired to present the full results in a satisfactory manner. However, some summary data on the nature of the complaints most frequently reported and on actions taken most frequently to seek redress or complain can be presented and this information seems interesting in its own right.

#### Specific Complaints

The Boulder instrument contained lists of specific complaints which the respondent could relate to the one item in each of the three sections which he had designated as "the least satisfactory" of all items in that section. The respondents were allowed to check as many of these complaints as they felt were applicable and multiple responses were common. The wording of the 15 different complaint statements for the nondurable products section and the percentages of the 275 respondents checking them are given in Table 1.

TABLE 1  
PERCENT OF 275 RESPONDENTS REPORTING SPECIFIC COMPLAINTS ABOUT THE LEAST SATISFACTORY NONDURABLE PRODUCT

(1) The quality of the ingredients and/or workmanship was inferior.	69.1%
(2) The product was not like the advertisement claimed it would be.	36.7
(3) The quantity was less than it was supposed to be.	32.4
(4) The package and/or the information on it was misleading.	28.7
(5) The product was spoiled or damaged.	17.5
(6) The product is unsafe.	16.7
(7) The product is harmful to the environment.	15.3
(8) I learned after buying that the "special discount price" was as high or higher than the regular price of other sellers.	11.3
(9) The instructions for using and/or taking care of the product were unclear or incomplete.	11.3
(10) The seller would not replace a spoiled/damaged item.	8.0
(11) An advertised "special" was out of stock when I went to the store to buy it.	7.3
(12) I was tricked into buying the product by a salesperson.	3.6
(13) I was charged a higher price than the one that was advertised and/or posted in the store.	2.9
(14) An item which was delivered was different than the one I bought.	1.8
(15) The credit terms were misrepresented to me.	1.1

The list of specific complaints available for respondents to relate to the least satisfactory item in the durable products section contained 17 items. The wording of each complaint and the percentages of the 275 respondents checking it is shown in Table 2.

The list of specific complaints which the respondents could relate to the least satisfactory category of services and intangible products contained 18 items. The wording of the statements and the percentages of the 275 respondents who checked them are given in Table 3.

It is interesting to note that while there is some variation over the types of complaints, there are definite patterns across the three categories. The most frequent responses in all three of the above tables were those directly related to the quality and performance of the product or service. Next in frequency were those responses having to do with unfulfilled claims or

advertisements for the product. Complaints relating to price and credit terms were relatively infrequent in the Boulder study.

- (16) Credit terms were misrepresented to me. 2.8
- (17) I was harrassed by bill collectors. 2.2
- (18) My insurance policy or service contract was cancelled without justification. 1.8

TABLE 2  
PERCENT OF 275 RESPONDENTS REPORTING SPECIFIC COMPLAINTS ABOUT THE LEAST SATISFACTORY DURABLE PRODUCT

(1) The product never performed as well as I was told that it would when I bought it.	61.8%
(2) The quality of materials was inferior.	54.2
(3) The quality of workmanship was inferior.	49.1
(4) The product has undesirable features that I was not told about when I bought it.	28.4
(5) The product was not all the advertisement claimed it would be.	26.9
(6) The warranty did not cover all of the things that when wrong.	24.0
(7) The dealer did not back up the warranty.	18.9
(8) The cost of use/operation is higher than I was led to believe.	15.3
(9) The product was damaged when delivered.	14.9
(10) The product is unsafe.	14.2
(11) The instructions for using and taking care of the product were incomplete or impossible to read.	10.9
(12) The manufacturer did not back up the warranty.	8.7
(13) The product is harmful to the environment.	8.4
(14) The price that was charged to me was higher than what I had agreed to pay.	6.9
(15) I was tricked by a salesman into buying a more expensive model/type than I needed.	6.9
(16) The item that was delivered was different than the one I bought.	4.4
(17) The credit terms were misrepresented to me.	3.3

TABLE 3  
PERCENT OF 275 RESPONDENTS REPORTING SPECIFIC COMPLAINTS ABOUT THE LEAST SATISFACTORY SERVICE OR INTANGIBLE PRODUCT

(1) The service was provided in a careless, unprofessional manner.	49.5%
(2) The service was not completed in the agreed time.	30.2
(3) I was charged for work that was not done.	26.2
(4) The quality of materials, parts, or medicines was inferior.	21.1
(5) Results fell far short of those claimed in ads.	20.4
(6) The fee was much higher than the amount agreed upon in advance.	17.8
(7) Performance and/or appearance was worse after the "repairs" than before.	17.5
(8) Professional services were rendered in an incompetent manner with very harmful results.	15.3
(9) I was charged for parts that were not furnished.	10.2
(10) Unauthorized repairs were made.	9.1
(11) The professional advice I paid for was incorrect and caused me substantial losses.	-7.3
(12) I was tricked by a salesman into buying services, insurance, or other intangibles.	6.2
(13) The premises and/or items that were rented to me were not as represented.	5.8
(14) The company refused to pay a valid claim.	4.7
(15) A professional confidence was violated.	3.6

Specific Actions Taken

Following the list of specific complaints relating to the least satisfactory item in each major section of the questionnaire was a list of possible complaint actions including the response "I didn't do anything at all." Respondents were asked to check off any of the actions they had personally taken. The wording of the statements and the percentages of the respondents checking each of them is given in Table 4.

TABLE 4  
PERCENT OF 275 RESPONDENTS WHO REPORTED TAKING SPECIFIC ACTIONS WITH RESPECT TO THE PRODUCT OR SERVICE REPORTED TO BE THE LEAST SATISFACTORY

Action	Durables	Non-durables	Services
(1) I didn't do anything at all . . . . .	22.9%	26.5%	16.0%
(2) I resolved to never purchase/use that product again . . . . .	49.1	54.5	39.3
(3) I told my friends of my experience and urged them to avoid the product. . . . .	49.8	45.1	47.3
(4) I returned the merchandise for a replacement/refund . . . . .	21.8	24.7	13.1
(5) I contacted the store to complain. . . . .	34.5	21.5	34.5
(6) I contacted the manufacturer to complain . . . . .	16.4	5.4	9.1
(7) I contacted the Better Business Bureau. . . . .	6.9	5.4	5.8
(8) I contacted my lawyer about possible legal action . . . . .	3.6	1.1	6.2
(9) I contacted a local or state official . . . . .	5.4	2.5	6.9
(10) I contacted my congressman or senator . . . . .	.7	1.1	1.8
(11) I contacted the officer of Consumer Affairs in Washington, D.C. . . . .	1.1	1.1	.4
(12) I contacted a consumer advocate to suggest an investigation (someone like Ralph Nader). . . . .	1.4	1.4	.4
(13) I contacted a federal regulatory agency (like the Federal Trade Commission) . . . . .	1.1	.4	2.2

It can be seen that a fairly substantial percentage of the respondents indicated that they did not do anything; however, the most common responses were those indicating that the respondent decided to boycott the seller and/or to urge friends to avoid the product or brand. Emphasis on the specific actions appears to vary considerably over the three types of products and services.

The data presented above were obtained in a single locality which is not highly typical of the national population. However, they are useful in illustrating some of the kinds of data which can be obtained in a

national study utilizing a large representative sample. It seems clear that such data along with additional data on levels of satisfaction, attitudes, and demographic data on the respondents can be of considerable value to public policy decision makers, business managers, and behavioral scientists.

#### Suggestions for Future Research

At this particular point in time it seems reasonable to place emphasis on descriptive research which can provide badly needed factual information on what consumers like and don't like and how they behave when they are dissatisfied with the products and services they purchase and consume. The data now available are clearly inadequate. While survey research is not without its problems, it seems that most promising way to obtain the comprehensive data needed now to evaluate consumer concerns, the response of business and government to them, and to assess the quantitative importance of various consumer problem areas.

Although it seems best to give priority to basically descriptive data at the present time, this does not mean that research aimed at explanation and theory building should be neglected. The goal of developing better conceptual structures for understanding and measuring consumer satisfaction should also be pursued. The development of better theoretical tools will be facilitated by the availability of better data and this in turn will allow the refinement and improvement of data collection methods. As the quality of factual data and our ability to interpret it are both increased, we will be better able to identify those areas in which the opportunities for dramatic improvements are greatest and those where the potential gains are limited.

While perhaps of a less fundamental nature than the measurement of consumer satisfaction and dissatisfaction, efforts to develop a useful "theory of consumer complaining behavior" should also go forward. As discussed above, the complaining behavior of consumers is not simply a matter of the perceived satisfaction or dissatisfaction with a product or service. It involves many other factors such as the personality, life style, and "informedness" of the individual as well as various situational factors related to the nature of the product itself plus the time, place, and circumstances of purchase and use. A better conceptual structure for evaluating consumer complaining behavior can be very helpful in understanding both the new data being gathered by survey research and the traditional complaint data which has been most important in the past and will undoubtedly continue to be a useful source of information in the future.

Hopefully, the development of better theoretical tools and improved measurement methods over the next few years will result in a reduced need for consumer research and increased levels of consumer satisfaction. It surely seems worth a try.

#### References

Ralph L. Day and E. Laird Landon, Jr., "Survey Data on Consumer Complaints for Consumer Protection Policy Makers," Proceedings, 1975 Midwest AIDS Conference, Indianapolis, Indiana, 40-44.

Betty J. Diener, "Information and Redress: Consumer Needs and Company Responses," Marketing Science Institute Working Paper, June 1975.

William R. Thomas and F. Kelly Shuptrine, "The Consumer Complaint Process: Communication and Resolution," Business and Economic Review, 21(June 1975), 13-22.

Raymond C. Stokes, "Consumer Complaints and Consumer Dissatisfaction," The Food and Drug Law Institute, Phoenix, Arizona, April 1974.

Rex H. Warland, Robert O. Herrmann, and Jane Willits, "Dissatisfied Consumers: Who Gets Upset and What They Do About It," The Pennsylvania State University, November 1974.

PUBLIC DEMANDS ON BUSINESS:

A Research Frontier

M. Denney, University of Michigan

Abstract

There is a general recognition that research is useful in pinpointing particular consumer concerns and problems. However, until recently there has been very limited recognition of more long-ranging and fundamental research efforts designed to construct a picture of more generalized concerns and impressions. This latter information is especially useful for businesses trying to develop a long-range view of consumer interests. One such effort in this respect is the establishment of a corporate priorities research effort undertaken by Yankelovich, Skelly, and White, Inc.

Corporate priorities provide business management with systematic overtime analysis of key public issues and how these are likely to affect corporate practices and profits. The substantive concerns around which corporate priorities are organized include consumerism, environmentalism, employee welfare, and general social responsibility.

One of the fundamental objectives of corporate priorities, in addition to supplying businesses with the aforementioned information, is to provide an assessment of when particular concerns will reach a "maturation" point and take shape as public policy issues. In this paper I examine the concept and application of corporate priorities and their relevance and importance in formulating and influencing public policies in the consumer domain.

Traditionally, consumer research has been concerned with the identification and analysis of factors relevant to individual spending and saving behavior. Those conducting and paying attention to such research fall into two main camps: 1) Marketing strategists, concerned with selling their firms' products/services; and 2) Economic planners, attempting to anticipate basic patterns of consumer activity. To these we might add a third category of consumer researcher, the academic, whose interests and motives have usually been less 'applied' than for either of the above groups.

Considering the concrete concerns and fairly well-established traditions of consumer research, it is perhaps not too surprising to find only very modest attention being paid, even now, to a new arena of investigation -- one that stands relatively distinct from the customary focus of either marketing strategists or economic forecasters. This new arena of research centers on the recently emergent phenomenon of consumerism, but extends really to what may be more generally regarded as public demands on business.

In the mid-to-late 1960's, there emerged a new force on the business scene, a new-found self-awareness among the nation's citizens in their

roles as consumers. The so-called 'consumerism movement' -- in both its formally organized form (Nader's Raiders, Consumers' League, etc.) and in its more diffuse manifestations (spontaneous product boycotts, etc.) -- can be viewed as a particular outcropping of broader social trends which have made themselves felt over the past decade or so. These include the various social rights movements, campus unrest, and new value trends with respect to sexual mores, the work ethic, and other basic concerns. At its heart, the consumerism movement appears to reflect a significant step in a broad process of politicization among the American mass public.

For the business community in particular, the emergence of this new public self-awareness and politicization has created a whole new set of game rules which have only barely begun to be recognized and brought under systematic study. Daniel Yankelovich has spoken of this as the challenge of "Marketing in a Climate of Mistrust:"\*

Anyone who has been alive and breathing these past few years knows what sweeping changes have taken place in the business environment in this remarkably short period of time. Perhaps the most troublesome change involves a shift from a climate of public confidence in business to one of skepticism and mistrust.

By now, most of us are thoroughly familiar with trend data which document the plummeting stature of the business community in the eyes of the public at large. But, what has been less aptly demonstrated by all the dire statistics is what the real implications of the new public demands may be, both for business people and for those in the public policy arena, who are without doubt just as directly the targets of this aroused public mood.

As consumers achieve ever greater political clout, there has come an increasing sense of urgency on the part of business leadership to try to understand just how serious a 'threat'-- if that be an acceptable term -- various sorts of public sentiments may pose. No longer are conventional market data on product preferences, demography, and life-style characteristics capable of providing (if they ever were!) sufficient information for the concerned business leader. New issues are at stake: product safety, truth in advertising, wage-price controls, environmental impact, and the like. And, new audiences and authorities have been drawn into the fray. Indeed, the very

\* Paper presented at GMA's Executive Conference, The Greenbrier; West Virginia - June 19, 1973.

conception of marketing strategy has had to be revised to take into account not only the consuming public, but the impact of governmental regulatory activity, as well.

The thrust of the new public demands on business has been not just to attack directly through the marketplace, but of at least equal importance it has been to bring tremendous new pressures to bear on those in public policy roles -- pressures to establish greater control and regulation over the way business conducts itself. The arena of conflict has been drastically widened as a consequence, and policy-makers are more involved than at any previous time in trying to comprehend and hopefully respond to emerging demands, needs, and interests of the mass (consumer) public. From the business perspective, it is safe to say that the impact of regulatory policies is of substantially greater concern than direct action in the marketplace. Product boycotts, picketing, and other forms of direct protest may be unsettling, but they seldom have the same depth of impact that regulatory legislation and other policy measures carry.

The challenge to the business community has been only very slowly taken up by systematic research comparable to that associated with conventional marketing studies. Private industry is naturally hesitant to invest the sizeable sums necessary to develop this new research focus. Far too often, this means that those in important decisional roles are forced to gauge and make judgments about the new public demands with little more to go on than intuition and whatever insights may be gleaned from public opinion polls appearing in the newspaper and weekly news magazines. Sad to say, but this state of affairs seems to apply equally to those on the public policy side of the picture!

But, the challenge being faced belongs not solely to the business community, nor to policy-makers. We in the research community have perhaps the most direct responsibility, that of providing the tools and strategies which will be essential for gathering this new kind of social intelligence.

Accordingly, my purpose here -- in addition to drawing needed attention to this challenge -- is to present a brief account of one research effort which has gone some distance in investigating the new public demands on business. I am referring to a cost-shared service known as Corporate Priorities, which was developed by the marketing and social research firm of Yankelovich, Skelly and White, Inc. at the beginning of the 1970's. While this undertaking has been geared primarily to a corporate clientele, it should nevertheless offer some important and useful lessons as to how such research may be extended for use by other audiences -- including those in the public policy arena.

#### Study Objectives

The central purpose of Corporate Priorities (CP) is to provide business management with systematic, over-time analysis of key public issues and how these are likely to affect

corporate practices and profits. This departs from conventional consumer research by being less concerned with product preferences, purchase motivations, and other matters pertinent to market behavior. Rather, CP gives primary attention to those sentiments and beliefs which have implications for the larger environment in which business operates, and particularly where formal regulatory action may be at stake. In other words, one major facet of this research venture is to serve as a sort of 'early warning system' for business -- to call attention to areas of tension and "risk," and to point toward effective responses.

Substantive Concerns. There is, of course, an immensely wide range of subject matter which may be included under the general rubric of public demands on business. The first order of business in any such research is to establish a framework and boundaries on the kinds of information being sought. In this regard, CP is organized around four major categories of pressure on the corporation:

- 1) Consumerism (pricing, product labeling, quality control, etc.)
- 2) Environmentalism (pollution, energy useage, etc.)
- 3) Employee welfare (working conditions, wage negotiations, etc.)
- 4) General social responsibility (corporate philanthropy, international involvements, etc.)

Admittedly, these are exceedingly broad categories, with each encompassing a host of complex issues and their varied manifestations. But, there are obvious limitations on how much information can be obtained, even in as ambitious a research enterprise as this one. Extensive preliminary investigation was essential for setting up a more detailed substantive agenda for the CP project, and this included both direct field study and consultation with members of the business community who would be the ultimate audience for this research. It is a demanding task to strike a proper balance between the recognized concerns and questions of the research clientele, without sacrificing the fundamental mission of this type of undertaking -- which is to find out what the public issues and demands are, rather than relying on conventional wisdom.

The task is far more complex than simply trying to identify a list of public demands or issues. There is, for example, only limited value in knowing that people want an end to polluted air if we don't also find out something about whom they blame for the problem, how serious a threat they think it really poses, what cost trade-offs they are willing to make, as well as any number of other considerations. Moreover, the number and complexity of such contemporary issues as covered in this kind of study yield a wide range of public opinion-holding. One must attempt to distinguish between superficial or "non-attitude" responses versus more serious viewpoints. One may also wish to determine what levels of information and sources of influence are

associated with different issue stances. Clearly the task is monumental.

Target Publics: To complicate matters even further, it should be rather obvious that sentiments found among members of the general public reveal only a part of the picture. While it is important to be in touch with what the population at large is thinking and feeling, it is perhaps even more vital for grasping the emergence of public demands to give careful attention to what is on the minds of various special and elite publics. Thus, one of the key features of CP has been its multi-targeted data base. In addition to a representative sampling of the general public, in depth interviews are also conducted among cross-sections of the following groups:

- ...Congressional committees in both houses
- ...Top officials in federal and state regulatory agencies
- ...Key members of the executive branch at the federal and state level
- ...Leaders of activist organizations
- ...Private and institutional investors
- ...College students
- ...Opinion leaders in the mass media, education, organized labor.

Specifying such targets is the easy part, however. The real problem begins with trying to determine the most effective and feasible methods for deriving samples among them, and for carrying out the actual interviewing. Where elites are involved, it is by now well recognized that conventional survey procedures must be adjusted to meet a variety of considerations. It is highly recommended, for example, that specialized personnel be used who are suitably matched to the particular kinds of respondents involved. The Yankelovich organization has recruited and trained a special corps of retired business executives for just such purposes; plus, senior professional staff are also called upon extensively to help gather some of the more important and sensitive interviews among special publics.

Such elite samples also require more than the usual amount of advance preparation, such as in setting appointments, making last-minute changes in schedule, etc. This can prove especially critical to the overall success of the project due to the need to coordinate closely all phases of data gathering and processing -- including elite samples, the general public, and so on.

Setting the Time-Frame. In many respects, CP represents a unique application of a 'social indicators' research approach; and, as such there is a basic requirement for a regularized over-time tracking design. Actually, this is a more complex consideration than it might at first appear for at least two major reasons: 1) Demands emerge and achieve prominence at widely varying rates; and 2) The total project involves massive data management tasks and

corresponding lag in turn-around time. In order to bridge the gap between a ponderous array of information and the need for currency, CP has developed a dual time-frame approach which works as follows: Once each year a full and comprehensive survey is made of all target publics, and this serves as the basic tracking period. But, in addition, to keep pace with faster moving issues and also to help meet special client needs, regular quarterly update surveys are carried out among only the public at large. This has proven to be a highly flexible and cost-effective approach particularly where public reactions form or shift rapidly, as in the case of product shortages, price increases, and other events which have occurred with distressing frequency in recent years.

Quantification. To a large extent, the methodological problems faced by CP are like those in other survey research -- i.e., problems of devising effective questions, turning verbal responses into numerical values, etc. There is no point here in hashing over standard procedures, approaches, or caveats. However, an undertaking like CP does encounter several more unique problems, stemming primarily from its function as a social indicators service. Uppermost is the task of data synthesis in moving from a profusion of atomistic information to a reduced set of more coherent and manageable indicators. And, out of this comes a second major problem, that of maintaining over-time continuity of such indicators from one survey period to another.

While the problem of data synthesis is not new, it has certainly found its greatest application in social indicators research. I am referring here not so much to what is sometimes called "data reduction" -- i.e., the selecting out of effective items from ineffective, and hence dispensable ones -- but rather to the task of constructing the most valid and reliable indices possible from combinations of items. It is precisely this indexing function which gives the social indicators approach its uniqueness and advantage over ordinary opinion polls. A composite index, after all, is designed to provide greater statistical reliability than any single question; plus, the compositing process itself tends to foster a greater depth of understanding as to the nature and limitations of the measures one is working with. This latter point cannot be too heavily emphasized, since the meaning and meaningfulness of indicators relied upon in this kind of research must be as thoroughly investigated as possible.

Aside from technical considerations, the need for data synthesis also derives from the kind of audience being addressed by such research. Corporate managers should not have to be research professionals in order to make sense out of CP data. Synthesized indicators can provide a valuable stepping-stone for turning data into findings -- thereby enhancing both the quality of analysis and the clarity of its presentation. This would be equally the case if one's audience were legislators, public administrators, or whomever.

Care must be taken, of course, to assure high quality in the indicators that are developed; and, for this a variety of statistical aids are

now widely available (factor analysis, clustering techniques, multidimensional scaling, etc.). But quite often researchers have hesitated to call upon such procedures because of their relative sophistication and a fear that clients would find the results obscure or somehow untrustworthy. However, my own experience indicates that this concern is largely ill-founded, with much depending on how one goes about incorporating such sophistication into the analysis. Business people want to know how the data speak to their particular problems and needs, without having to wade through a morass of statistical mumbo-jumbo. It is important, therefore, to forego the academically inherited penchant for formal authentication, and to take a more positive posture in the use of synthesized indicators.

This does not mean, of course, that statistical fine-points may be ignored behind the scenes, as it were; and, this brings us to another aspect of the use of composite indicators. This is the problem of maintaining continued validity and reliability of such measures over time. The classic issue in this regard is whether measurement structure remains the same over successive tracking periods. Or, stated more simply, whether the same sets of items still "fall together" or otherwise satisfy the criteria which initially justified their being combined into specific indices. As events progress, it is only natural that attitudes and issues will evolve and change their configurations. New demands arise, old ones fade. Broad issues take on different specific manifestations, and so on. Only a few years ago, for example, a question about the desirability of selling grain to the Soviets would have evoked responses based primarily on anti-communism sentiments, and thus perhaps qualifying as part of a "trade-with-communist-nations" indicator. However, the mounting pressures of inflation plus the somewhat infamous wheat deal with the U.S.S.R. two years ago have sharply altered the situation. Now, the meaning of such a question is undoubtedly more aligned with the extent to which individuals feel that foreign grain sales would create (more) unwelcome pressures on food prices in this country.

While the re-structuring of public demands and issues can be troublesome for measurement continuity, it is also one of the more important matters upon which a well-conceived research program of this sort can provide information. In this sense, then, the ongoing validation of indicators is more than just a job of fine-tuning or statistical finesse. It can offer extremely useful insights on the substantive side of things, as well.

#### The Policy Process: Putting It All Together

The discussion so far has centered mainly on certain methodological problems in doing research into public demands on business. In these concluding remarks, however, I want to shift attention toward the theoretic level and specifically to the question of how information gathered in this kind of endeavor can be integrated into a meaningful picture of the policy process as it affects business. It is sufficiently challenging to try to identify and track salient public issues from their in-

ipient stages all the way to their being converted into formal regulatory policy. But, there remains a further and ultimately more critical challenge -- that of unraveling the paths and mechanisms through which the process of public policy formation takes place.

Ideally, we would like to have a theoretically derived model of the policy process, and one from which predictive analysis could be waged. But then, who wouldn't? Lamentably, there is far too much complexity, as well as plain enigma, surrounding the translation of public sentiments into public policy to allow an easy rendering into "model" form. At present, scholars have had only modest success in depicting the inner workings of subcomponents of the policy process (Congress, political parties, elections, etc.), and have barely scratched the surface at higher levels of analysis.

Short of an actual theory or model of the policy process, however, it is feasible to establish a working paradigm. In the case of CP, the paradigm that has been adopted is, for obvious reasons, built closely around the central concerns of the business clientele. A key unifying element that comes from this is the concept of "maturation" of demands. This refers to the succession of stages through which public demands emerge, gain momentum, attract coalitions, and are carried (or fail to be!) into formal regulation. A primary objective of analysis, then, is to determine the degree of maturity of specific demands and hopefully provide some indication of future developments.

In terms of research strategy, this objective is reflected in two major facets of CP's design discussed earlier: 1) tracking over time; and 2) multiple target publics. Each of these elements taken alone offers important predictive advantages, but in conjunction they present a potentially powerful grasp of policy formation. There is not only the opportunity to view trend lines within the general public and within special and elite publics, but also to follow the patterns through which demands move from one major public to another within the broad policy process. This, it seems to me, is a large accomplishment and a valuable step toward realizing a more theoretically elaborate model of policy formation sometime in the future.

For the present, the goals are practical and limited, the approach heavily empirical. Gradually, however, as the CP project accumulates findings, experience is gained in interpretation. With this should come a corresponding growth in understanding of how public policy evolves. Of course, this growth of understanding will be very slow without more research along these lines being undertaken. There is a curious irony in the fact that research into public demands on business is itself at least partially an outgrowth of an important theme among such demands -- that for greater corporate responsiveness to the public good.

## ANSWER FROM EXPERIMENTAL PSYCHOLOGY

J. Edward Russo, Psychology Department  
University of California, San Diego

## Abstract

The work of Jacoby and Small (1975) on misleading drug advertising is extended. The recommended procedure includes an objective measure of misleadingness (misprescriptions), an amended, non-misleading control advertisement, and a statistical test. Several more sophisticated procedures are offered. The analysis separates the information providing role of the social scientist from the value judgement role of the policy maker.

Many public policy makers must answer the question: is a given product harmful to the public? For example, judgements must be made about the safety of products from toys to automobiles. If harm is judged to exist, the policy maker must select an action: a warning, withdrawal of the product, or restitution for damages.

The problem facing the policy maker may be partitioned into two parts. First, he must establish that the product causes harm. Second, corrective action must be based on the seriousness of the harm. It is essential to distinguish these two questions. Social scientists, serving as consultants to policy makers, are qualified to address only the first issue, the existence of harm to the consumer. The value judgements required to answer the second question, seriousness and corrective action, are the proper role of the elected or appointed policy maker. In keeping with this distinction the present paper is concerned only with deciding the existence of harm. One specific situation is considered, but policy makers should be able to generalize the proposed technique to their own policy setting.

## Misleading Advertising

The problem to be analyzed here is the identification of misleading advertisements. Jacoby and Small (1975) discuss this problem for the specific case of the advertising of drugs, especially to physicians in professional journals. This paper is an extension of their analysis of the problem.

Jacoby and Small point out a fundamental difference between the approaches of the Federal Trade Commission and of the Food and Drug Administration to the problem of misleading advertising. The latter attempts to demonstrate the intention to deceive, often relying on the testimony of experts and the adversary procedure that characterizes the judicial process. In contrast, the FDA focusses on the effects of improper drug advertising, i.e., whether a physician reading the ad is misled. This latter approach has two primary characteristics. It is consumer oriented, and it utilizes empirical findings. It is consumer oriented because the central concern is whether harm was done to the consumer, such as by a physician's improperly prescribing the advertised drug. In contrast, the FTC's strategy of demonstrating the intent to deceive ignores the effect on the consumer. The FDA's approach can utilize empirical knowledge because the question of misleadingness can be answered empirically. In contrast, the question of deceit is determined by traditional legal procedures.

Jacoby and Small (1975) proposed the following empirical procedure for determining misleadingness. First, devise a measure of misleadingness that is minimally susceptible to bias. For example, rather than ask physicians to judge directly whether an advertisement is misleading, ask whether they would be likely to prescribe the drug for some inappropriate ailment. A positive reply would indicate that the advertisement has misled the physician. Second, because prior attitudes toward advertising or drug manufacturers may influence the tendency to perceive misleadingness (Haefner, 1972; Koffman, 1964), select a representative panel of physicians. That is, let misleadingness be estimated from a group representative of those likely to read and use the advertisement in question. Third, employ a control advertisement, which would be shown to a second representative panel of physicians. The control advertisement would be modified to eliminate the tendency to mislead. For example, the questionable advertisement could be amended by excising a claim or inserting a warning.

Based on such a procedure, Jacoby and Small propose the n% criterion. If more than n% of the physicians are misled, then the advertisement is judged to be misleading. As the authors acknowledge, this is only a partial solution since a criterion value of n% must still be chosen. The real problem is that a single value of n may not be equally suitable in all situations. For example, if the advertised drug were very powerful the rate of misprescriptions might be considerably lower than that of a less potent drug. Similarly, the chance of being misled may depend upon how much is known either about this class of drugs (e.g., side effects), or about the ailment in question. For all these reasons, variation in the n% criterion must be expected. The task is, in the face of this variation, to find a procedure for determining that percentage of misled physicians which implies that the advertisement is truly misleading.

In the next three sections, three different procedures for establishing misleadingness are discussed. Each new procedure increases the level of sophistication, at some cost in data collection and in transparency of the results to the policy maker.

## Procedure 1

This procedure relies on a control advertisement and a yes/no determination of whether a physician has been misled. The control advertisement would be as similar as possible to the purportedly misleading advertisement, except for the removal of the misleading aspects. To determine whether each physician tested was misled by exposure to either the real or control advertisement, a simple question could be asked. For example, "Would you prescribe this drug for Ailment X?" (Ailment X would be selected as that ailment for which misprescription would be most harmful and most likely.) Responses to such a question would yield the following data:  $P_M$  and  $P_C$ , the proportions of physicians misled by the real and control advertisements, respectively. Note that  $P_M$  is the n% of Jacoby and Small.



Given both  $P_M$  and  $P_C$ , standard statistical techniques for hypothesis testing can be employed.  $P_C$  is an estimate of the misleadingness that is due to chance alone, i.e., to all factors other than the misleadingness supposedly in the real advertisement. Any excess of  $P_M$  over  $P_C$  should be caused only by the misleadingness in the questionable advertisement. Using a standard test for the equality of two proportions, we can test whether the target advertisement is misleading. For example, let  $P_M = .15$  and  $P_C = .02$ , where both proportions are based on the responses of 200 physicians. It can be shown that if we assume that there is no misleadingness (the null hypothesis), then the probability ( $p$ ) of observing as large (or larger) a difference between  $P_M$  and  $P_C$  is .00029. Using the standard analytical procedures of statistical hypothesis testing, we reject the null hypothesis and conclude that the target advertisement must have misled the physicians who read it.

Of course, it is the nature of statistical inference that one admits to less than absolute certainty in one's conclusion. Thus, there is a preset level of significance ( $\alpha$ ), such that if the observed value of  $p$  is less than  $\alpha$ , one concludes that misleadingness exists. The  $\alpha$ -level must still be set by judgement (that of the policy maker, not of the social scientist). Unlike the  $n\%$  criterion, however, choosing an  $\alpha$ -level is a well understood aspect of statistical practice.

The essence of this first procedure is the use of a control group to estimate a base rate. The reasoning involved is relatively simple and should be comprehensible to both policy makers and their constituents. For a similar application of the same techniques to a problem in criminal justice, see Russo (1975). Finally, note that the data,  $P_M$  and  $P_C$ , are also relatively easy to collect. One must only devise an instrument for a Yes/No determination of misleadingness.

#### Procedure 2

Procedure 2 differs from its predecessor in one respect only. It uses a numerical measure of misleadingness. I will not attempt to describe in detail an instrument for generating such a numerical rating. It might be based on a total misprescription score over several ailments. Even simpler, instead of asking each physician for a Yes/No prescription decision, one might ask for the percentage of cases of Ailment X for which he believes the advertised drug should be prescribed. The same basic question is being used, but the response requires more information, a numerical estimate of prescription rate rather than only a Yes or No.

The availability of a numerical measure permits the use of more powerful statistical tests, notably the t-test. Whether or not this advantage outweighs the increased effort needed to collect the numerical estimates of misleadingness is a question best answered by experience. Only by pretesting an instrument on a sample of physicians can we determine whether the more valid responses will be given to the Yes/No or to the numerical versions of the question.

#### Procedure 3

One of the most bothersome aspects of the empirical determination of misleadingness is the existence of a prior tendency toward judging an advertisement as misleading. As noted earlier, attitudes toward adver-

tising, drug companies and so forth, may contribute significantly toward the tendency to judge any advertisement as misleading. Jacoby and Small (1975) handle this problem in two ways. First, misleadingness is determined from questions that are as objective as possible (e.g., specific prescription behavior) and not from questions that make it easy for the bias to enter. Second, both experimental and control groups are selected such that any prior bias toward misleadingness will occur equally in both groups. The combination of these procedures should be adequate to insure the validity of either Procedure 1 or 2. Nonetheless, if additional protection against the effects of this bias is desired, a more sophisticated procedure can be employed.

The proposed procedures rely on a measure of the prior bias for each physician questioned. For example, one might ask, "What percentage of drug advertisements (in medical journals) do you consider to be misleading?" The responses to such a question can be used in several ways. First, one could check that the experimental and control groups were really balanced for prior attitude by performing a t-test on these data. Second, one could check the objectivity of the question that is used to determine the misleadingness of the target advertisement by testing for a relation between the measure of prior bias and the measure of misleadingness. If there is no relation, then the misleadingness question is truly objective. If there is a relation, then a matched t-test could be performed, with the physicians in the two groups matched according to their prior bias levels. Such a matching procedure removes the effect of prior bias on the misleadingness scores (similar to a repeated measures or within subjects design in the analysis of variance). This also enables a more sensitive detection of the existence of misleadingness.

#### Other Procedures

The preceding procedures are not exhaustive. The theory of signal detectability (Green & Swets, 1966; Coombs, Dawes & Tversky, 1970, Chapter 6) might be applied to this problem. The decision criterion would be stated in terms of  $d'$  rather than  $Z$  (or  $p$ ). Because the signal detectability paradigm provides a systematic method for deciding if the specific questions being asked are biased for or against misleadingness, the prior disposition toward perceiving misleadingness could be investigated in more detail. One possibility is a multiple regression analysis for identifying physician characteristics (age, income, specialty) that contribute to a prior bias. The point of mentioning these techniques is not to recommend them at the present time, but rather to illustrate the range of empirical information that can be made available to the policy maker.

#### Implication for Policy Formation

Two aspects of the procedures that have been presented are critical for the policy maker. First, all final judgements remain in the control of the policy maker. All he gets from the social scientist consultant is information that will help him make those judgements. For example, it can be determined (at specified levels of certainty) that some misleadingness is or is not present in the target advertisement. The remedy adopted, however, is still up to the policy maker. If he feels that the harm caused is slight, simple withdrawal of the offending advertisement may be sufficient. If the harm is major, however, then corrective advertising could be required. The available options should not be reduced by the use of a sound, empirically based analysis of misleadingness.

The second important aspect of the proposed procedure is that the policy maker has several levels of sophistication to choose from. This advantage is often undervalued. The policy maker's freedom to choose any one of several techniques may be restricted by factors beyond his control. For example, in some situations it may not be possible to obtain a numerical measure of misleadingness. In such cases, only Procedure 1 is available. Alternately, the policy maker's constituency, such as a higher policy body or the drug companies and their advertisers, may not comprehend and accept a more sophisticated procedure. The point here is that more than analytical constraints may affect the acceptability of the empirically based analysis that a policy maker can use. By offering several levels of sophistication, the present proposal does not confront the policy maker with a "take it or leave it" situation. Rather he can choose the level of information that will best serve a variety of policy making considerations.

#### References

Clyde H. Coombs, Robyn M. Dawes and Amos Tversky, Mathematical Psychology: An Elementary Introduction (Englewood Cliffs, New Jersey: Prentice-Hall, 1970).

David M. Green and John A. Swets, Signal Detection Theory and Psychophysics (New York: Wiley, 1966).

James E. Haefner, "The Legal Versus the Behavioral Meaning of Deception," in M. Venkatesan (Ed.), Proceedings, Third Annual Conference for Consumer Research, 1972, 356-360.

Jacob Jacoby and Constance Small, "Deceptive and Misleading Advertising: The Contrasting Approaches of the FTC and the FDA," Journal of Marketing, 1975, in press.

E. John Kottman, "A Semantic Evaluation of Misleading Advertising," Journal of Communication, 14(September, 1964), 151-156.

J. Edward Russo, "A Statistical Analysis of Wiretap Evidence," unpublished manuscript, University of California, San Diego, 1975.

THE FORMATION OF CONSUMER POLICIES IN THE  
PUBLIC DOMAIN: A CONCEPTUAL OVERVIEW

Saul Barry Wax, Massachusetts Consumers' Council

Abstract

One intent of the panel was to suggest that the public and private sectors should be thought of as parts of a whole, and not distinct entities unto themselves. The result of this conceptual schema is to recognize the formation of consumer policies as a more consensual and less adversarial activity. Another objective of the panel was to clarify the ways in which social scientists could have input into the policy-making process. Both of the objectives were met. In addition, it is suggested that the linkage of the public and private sectors provides a fertile opportunity for social scientists to utilize their professional skills and help influence and direct the creation and implementation of consumer public policies.

There are two major issues which I hoped this panel would illuminate: the increasing linkage between the public and private sectors, and that social scientists could and should have a larger role in the explication and resolution of public policy issues. I am pleased to report that my objectives were realized, due in large part to the participants in this panel, and their respective presentations.

As a policy analyst working for a public agency which attempts to influence, shape and direct public policy, especially in the consumer field, it has become very apparent in recent years that the conventional conceptual distinctions between the public (governmental) and private (business) domains is increasingly less accurate as well as less useful for understanding the formulation of consumer policies. There are several reasons for this change and I will briefly discuss them here.

Insofar as accuracy is concerned, a sketchy replay of immediate past events shows that the public and private sectors are not distinct. First, we have the obvious movement of industry personnel to the governmental sector. Both the Departments of Commerce and Agriculture are frequently referred to as "service agencies" for their respective clientel groups. And officials from affected or concerned firms frequently take governmental jobs, a tour of duty so to speak, and then return to their companies. The recent grain sales to the Soviet Union starkly illuminated the movement of government officials to the grain companies.

The distinction between the public and private sectors is not really accurate for a second reason: the problems which occur in the private sector are often resolved by the public sector on terms favorable to the private sector entity. Indeed, if the two sectors were, in any sense, adversaries, then one might expect that certain concessions and prohibitions might be extracted from the private sector. But the Lockheed case suggests otherwise. And the New York City fiscal situation is equally revealing. The operant concern about New York expressed by public officials in Washington is that the banks not be injured nor the stockholders lose their investment. Federal aid is thus available to avert that possibility.

In yet a third respect the distinction between the two domains is not accurate. Not only when responding to crisis situations, e.g., Lockheed, New York City, is

the symbolic relationship between the public and private sectors apparent, but on a day-to-day basis the linkage is equally if not more present. At least five Cabinet Secretaries (Federal) are directly responsible for funding and maintaining large segments of the nominally "private" sector.

Department of Defense	Defense or defense related industries
Department of Housing and Urban Development	Construction and related business
Department of Agriculture	Agribusiness firms and producers
Department of Commerce	Business firms (general) with concern with international trade
Department of Health, Education and Welfare	Supports social welfare programs and firms (including schools)

In effect, normal operating procedure includes the ongoing support of numerous private sector firms. With so many synapses between the two, it might be preferable to recognize that the "public" and "private" sectors are really segments of the same organism, and if any part of the whole becomes ill, it could damage the entire system. Thus, as a routine, steps are taken to try and insure that "illnesses" will not occur.

A final reason for the inaccuracy of the distinction is that arguing that there are public and private sectors involves, at least implicitly, recognizing the existence of public and private interests. Yet when the failure of a Lockheed or New York City bank threatens the welfare of a massive number of people, is it not in the public interest for the government to respond?

I am not arguing that there are no clear and distinguishable public versus private interests. Rather, I believe that in at least some instances even the concepts of "public" and "private" interests cannot accurately be affixed to particular outcomes.

In addition to being of questionable accuracy, the distinction between public and private sectors is not especially useful for increasing our understanding about the formation of consumer policies for two reasons. First, policies which affect the public emanate from the "public" as well as "private" domains. When policies come from the government, such as constructing an office building, they are called public and when they come from the private sector, as in the case of supermarket check-out equipment, the term "business policies" is affixed to them. But if they both affect the mass public, it might be more appropriate to classify a policy in terms of its impact and not its genesis. Thus, all segments of the society can affect and influence the consumer regardless through the policies they establish.

The second reason is that neither segment of the society will function well apart from or without the other: for example, the government needs Lockheed and the latter surely needs government. The Federal Reserve Bank

serves the banking industry. Indeed, it would be difficult to imagine any single segment of the society carrying on in vacuo. Consequently, the distinction between public and private becomes all the more arbitrary and not useful in understanding how consumer policies are formed.

As an alternative to the public/private sector distinction, I suggest that the society be conceived of as a systematic whole. Further, that consumer policies reflect the needs of the entire system. In effect, consumer policies are far less adversarial than accomodating. And it is here where the social scientist can come into his/her own.

Assuming that policies are created in the "cooperative" (for lack of a better term) framework described above, it is also arguable that an objective and clear explication of the components of a policy issue are necessary to its resolution. And since we're talking about policies which affect mass publics and thus society, then social scientists should be unusually well equipped to be of assistance. This situation is all the more probable with the increasing complexity of the issues and the concomitant growing demand for information.

Both the Denney and Russo papers provide keen insight into how social scientists can become involved in and influence the formation of public policies. Denney looks at social science information being fed into business concerns, while Russo takes the same task with respect to government. Together the papers serve to illuminate the kind of environment within social scientists may act, and the way in which they may have access and impact.

Note: the views expressed herein are solely those of the author and do not in any way reflect or represent the attitude, belief; or policy position of the Massachusetts Consumer's Council.

EXPERIMENTATION FOR PRETESTING PUBLIC HEALTH PROGRAMS:  
THE CASE OF THE ANTI-DRUG ABUSE CAMPAIGNS<sup>1</sup>

Michael L. Ray, Stanford University  
Scott Ward, Harvard University

In recent years, much attention has been devoted to "broadening the marketing concept," i.e., applying marketing concepts and techniques to problems beyond the traditional private enterprise sector. One such zone of application has been public health related problems. As health-related fields realized the necessity of effective communication with target audiences, social researchers in universities and consulting agencies were quick to respond to the resulting need for campaign evaluation research. Comparatively little has been done to illustrate to public health professionals the application of marketing concepts and measures to problems involved in pretesting. By "pretesting," we are referring to the activities of defining communication objectives, generating message alternatives, selecting the most effective messages and campaign strategies, and evaluating results.

This paper summarizes a research project which sought to develop pretesting methods for federally-funded anti-drug abuse media campaigns. The project consisted of two large-scale experiments in laboratory settings, and a field experiment utilizing a split-cable facility. The field experiment was designed to provide field comparison data to tentatively assess the validity of less natural laboratory predictions.

The project focused on selection and scheduling of commercials, i.e., the problem of selecting one finished television commercial among several alternatives designed for various audience segments, rather than the problem of developing procedures for "idea generation," or some other aspect of pretesting. The project used finished television commercials, but the procedures developed here could readily be adapted for use with other materials (e.g., storyboards), and the procedures could also be adapted for pretesting advertising for media other than TV.

#### Overall Project Objectives

In planning and conducting this research, we pursued the following general objectives.

1. To develop a pretesting technique that is reliable, valid, and sensitive to different audience effects produced by various advertisement alternatives.

To this end three different commercials -- "Walkout," "Big Brother," and "People" -- were tested among three different audience groups. Two large-scale laboratory experiments were conducted in order to test and refine measures and procedures, and a field experiment using a split-cable facility was conducted in order to validate these laboratory procedures.

2. To develop a practical pretesting method.

To this end the procedures we recommend could provide preliminary data within two weeks after receipt of test advertising; out-of-pocket costs would be approximately \$1,750.

3. To create a test environment which approximates "normal" television viewing, i.e., in which conditions are manipulated to reflect "real world" viewing conditions, and which can be expected on theoretical grounds to affect audience responses to advertising in important ways.

In order to accomplish this objective, subjects were exposed to the test advertising in naturalistic surroundings, and the ads were imbedded in a 17½ minute segment of programming. Additionally, we varied the level of distraction present during the pretest, and the nature and amount of competing messages during exposure.

4. To employ a range of dependent ("effects") variables which reflect long-term and short-term cognitive and behavioral effects of advertising, and which are consistent with advertising objectives.

To this end, we measured cognitive responses to the advertising which occurred during exposure, by asking subjects to recall what they thought about while viewing. From these questions, we found subjects frequently "counterargued" against the messages, and/or "connected," i.e., experienced a mental link between message content and self-experience. We also examined the following effects:

- recall of points made, sponsor, situation shown;
- changes in ranking of the "drug abuse problem";
- attitudes about anti-drug abuse;
- interest in receiving a booklet about anti-drug abuse;
- booklet reading; and
- stimulation of discussions about anti-drug abuse with family or friends.

#### Overview of Procedures

In pursuing our study objectives, we employed a technique which involved each respondent in the following series of events:

- a. Come to a central location
- b. Read "Television Violence-Humor Project," the cover story for the experiment.
- c. Complete sociodemographic and media exposure questionnaire.
- d. View a 17½ minute program on a closed-circuit television set. The program contains two exposures of test anti-drug abuse commercials as well as regular programming seen on television.
- e. Complete self-administered questionnaires with items relating to the cover story, as well as items eliciting information on recall of advertising, cognitive response, attitudes toward social problems, and interest in receiving

<sup>1</sup>This research was supported by the National Institute of Mental Health (NIMH-OC-72-156) and the Marketing Science Institute.

relevant booklets. Specific attention is not paid to drugs or anti-drug abuse advertising until the last questionnaire.

- f. Sometimes, following the above study procedures, drug booklets are sent to respondent's homes and follow-up telephone interviews are conducted to determine their reactions.

The technique developed in this project can be adapted to different types of respondents (e.g., children vs. adults), different test materials, viewing conditions, and measures employed.

Reliability and validity of the pretesting technique was assessed by examining consistency of results across three of the project studies -- two laboratory studies and one split-cable field experiment. The technique appears to work well because it simulates natural viewing and response processes, a simulation produced by conditions of distraction and attention developed in the first study. It should be pointed out, however, that the reliability and validity of the technique is limited to gross indications of the relative effectiveness of advertisements in various conditions. No pretesting technique can produce accurate absolute predictions, but this technique will offer excellent guidance for public health campaign planners.

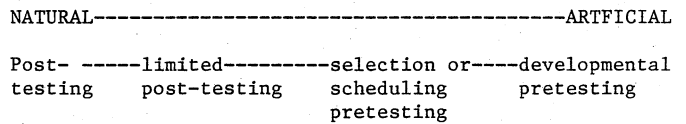
The split-cable field experiment showed that it is possible to produce meaningful effects with a heavy anti-drug abuse campaign in a short period of time (one month). A campaign with such a heavy emphasis could only be carried out in short, local efforts, but such a campaign seems well worthwhile given the frequent, local, epidemic character of the drug problem. The campaign would have promise for promoting an atmosphere of belief in which other anti-drug abuse activities could work effectively.

Three sets of dimensions are relevant in evaluating selection or scheduling pretesting procedures:

- design, measurement, and sampling dimensions;
- creative (or stimulus) research dimensions; and
- cost and payout dimensions.

For each set of dimensions, campaign planners must make decisions based on the naturalness or the artificiality of the pretest research setting. A completely natural test would involve post-testing, or running a complete communication campaign on the media and observing responses in terms of actual drug abuse. A completely artificial pretest would be the type referred to above as a developmental pretest, the kind almost always conducted in anti-drug abuse campaign development. Obviously, if there were unlimited funds and time, the most natural type of test would be desirable. It would be totally valid and would provide much valuable information for long-range development. On the other hand, less natural forms of pretesting are less expensive and have great short-term value because they allow for changes in campaigns before running them full-scale with possible damage.

The three types of pretesting discussed to this point can be arrayed on a dimension from natural to artificial, as shown below:



As can be seen, the selection/scheduling type of pretesting is arrayed near the middle of the natural to artificial dimension. Thus, this type of pretesting has the advantages of a relatively high degree of validity and precision which only a completely natural test can achieve, as well as the efficiency of the more artificial types of pretesting.

Design, Measurement and Sampling Dimensions

The specific dimensions involved in design, measurement and sampling considerations are arrayed on the natural to artificial scale in Figure 2.

FIGURE 2  
DESIGN, MEASUREMENT, SAMPLING DIMENSIONS

<u>DIMENSIONS</u>	<u>NATURAL TESTING</u> -----	<u>ARTIFICIAL PRETESTING</u>
DESIGN	In-home or school-----Familiar Quarters-----	Mobile unit-----"Downtown"
M E A S U R E M E N T	Natural, unobtrusive----Coupon-----	Choice in test-----Obtrusive measure measure (behavior) environment
	Repetition effect measured-----	-----One shot measures
	Lag effect measured-----	-----Immediate measure
	Respondent alone responds-----Self-administered with-----	-----Interviewer makes responses interviewer supervision
and	Individual or family----Individual-----	Family group-----Unusual group
S A M P L I N G	Representative-----Just random-----	Matched-----Accidental sample random sample
	Complete experi- -----Control of time-----	Possibility of-----No control or com- mental control of observation correlational parison possibility study

As the figure illustrates, campaign planners must make decisions concerning research design: should the research be conducted in home, or in an artificial research facility? Measurement and sampling decisions must also be made: should the measure be unobtrusive (i.e., totally natural); should repetition effects be taken into account, should the sample be representative and random, or can an "accidental" sample be tolerated? Perhaps most important, should the design and measurement plan be based on total experimental control (i.e., of temporal order in presentation of test stimuli), or

should the research be "one-shot," with no comparison or control possibility?

Creative (or Stimulus) Research Dimensions

A second set of dimensions involves decisions about what, specifically, subjects or respondents will respond to in the pretest. The pretest dimensions involved in subjects' responses to various types of test ads are arrayed on the natural to artificial scale below.

FIGURE 3

CREATIVE (OR STIMULUS) RESEARCH DIMENSIONS

	<u>NATURAL TESTING</u> -----	-----	<u>ARTIFICIAL PRETESTING</u>
S	Campaigns-----	Representative ads	-----
T	Finished-----	Rough film-----	Photo on film-----
I	Commercials		Art skills-----
			Storyboards-----
M	Natural sur- -----	Program-----	Program-----
U	rounding material		Other messages-----
			Surrounding material
L	(Magazine)-----	(Skeleton issue)-----	(Portfolio)-----
			(Ads alone)
I	Ordinary -----	On TV, reg- -----	On TV, spe- -----
	presentation	ular programs	cial programs
			Closed-----
			Film into-----
			TV like screen
			Handy presentation

Key decisions concern whether individuals will view complete, finished campaigns, or simply rough cuts or storyboards. Another important decision concerns whether people view test ads in isolation, or in the context of other messages. If respondents simply view a test ad on a tear sheet, this is obviously more artificial than viewing the ad in the context of a dummy magazine. In the case of television messages, it is obviously more "natural" to view test ads via a television received rather than via a movie screen; and it is also more natural to include test messages in the context of programming and other messages, rather than in isolation.

and interest in receiving more detailed information (a booklet).

- The "Big Brother" commercial was intended to stimulate older adolescents to warn younger siblings about the danger of drug abuse.
- Compared with the above, the "People" commercial was more amorphous in its objectives. It contained "face shots" of various individuals (including celebrities) remarking about the drug abuse problem, and concluded with an appeal to the viewer to send for a booklet about it.

Cost and Payout Dimensions

Finally, the campaign planner must make decisions based on costs, and on how quickly the results must "payout". The pretest costs, time, and payout dimensions involved in planning media campaigns are arrayed on the natural to artificial scale below.

For selecting independent variables other than the ads (e.g., viewer attitudes, distractions, or competing messages which resulted in the commercial having greater or lesser impact on viewers), the following criteria were used:

1. Independent variables should induce conditions reflecting "real world" television viewing.
2. They should produce differences in effects.
3. They should permit testing of general expectations for the data, on theoretical grounds.

FIGURE 4

COSTS AND PAYOUT DIMENSIONS

<u>NATURAL</u> -----	-----	<u>ARTIFICIAL</u>
High cost-----		Low cost
More time-----		Less time
Little short-term payout-----		More short-term payout
Much long-term payout-----		Little long-term payout

For selecting dependent (effects) variables (e.g., cognitive responses to ads, degree of recall, booklet interest, ranking of drug abuse problem), the following criteria were used.

1. Dependent variables should be valid, reliable, sensitive and practical.
2. Multiple measures should be employed to illustrate the range of effects.
3. The measures should reflect people's thoughts while watching.
4. The measures should be consistent with advertising objectives -- namely, to measure both cognitive response and behavioral reactions.

Initial Study

Research Procedures Used in Study One

In the first study, conducted in a laboratory setting, tests were run on three 30 second anti-drug abuse ads, ostensibly aimed at three different audience groups:

- The "Walkout" commercial, aimed primarily at parents of adolescents and preadolescents, was intended to stimulate concern about drug abuse,

## Independent Variables

### • Level of Distraction

A key variable employed in the first study was the level of distraction present during exposure. As television viewing normally occurs under distracting conditions--people talking, etc.--the theoretical literature suggests that such distraction could be expected to affect responses to ad stimuli. In early persuasion research, investigators suggested that subjects generate counter-arguments while attending to messages opposed to their position, and that distraction while subjects attend to persuasive messages might inhibit counterargument production, thereby facilitating attitude change. However, distraction cannot be so great as to eliminate message comprehension, and the topic must be a salient one for subjects (see Festinger and Maccoby, 1964; Baron, Brown, and Miller, 1973; Roberts and Maccoby, 1974).

In the present study, we expected varying degrees of commitment to the position that drug abuse is a serious problem, that it is harmful, etc. Adults should be most committed to the position advocated in the ads, while teenagers and preteenagers may be less committed. In any case, we expected varying levels of distraction during viewing to alter the amount and kinds of thoughts these different audience members experienced during viewing.

### • Competing Messages

A second independent variable employed in the first study was amount and kind of competing messages--i.e., other public service ads, advertising for over-the-counter drug products, and a control condition (no competing ads).

Dependent Variables. In most pretesting research, responses to ads are obtained following exposure. Such viewer reactions may take the form of several cognitive responses, such as counterarguments against the message, the situation portrayed, etc., or their reactions may be in the form of thoughts which essentially link message content and personal experience ("connections"). These responses most often concern reactions at the time the question is asked, and considerable bias can enter into the question/response procedures.

Consequently, we presented an open-ended question, asking respondents to describe "what they thought about" while watching the ads. On the basis of previous research, we expected most reactions to be in the form of counterarguing, and "connections"--which follows from research and theoretical notions of Herbert Krugman (1968a, 1968b). Often investigators have posed different cognitive responses in persuasion research, e.g., "rehearsal" of message arguments (Kelman, 1953); counterarguing support argument, source derogation, and curiosity (Wright, 1972).

Based on this research, the objectives of the messages, and our knowledge of the orientations of the subject populations from previous research on drug attitudes and usage (Johnston, 1973; National Commission on Marijuana and Drug Abuse, 1972), we specified the following general expectations for our data.

1. Since adults can be expected to hold attitudes consistent with the anti-drug abuse messages, they should counterargue less, and experience more connections than teenagers. However, since the anti-drug abuse messages contain unpleasant, perhaps threatening information, adults should

be particularly sensitive to level of distraction due to the high degree of cognitive effort required to counterargue or connect.

2. Many senior-high-school students can be expected to be opposed to the anti-drug abuse messages, and to possess much information and extensive predispositions about drug abuse. Therefore, they should counterargue more than other groups, but increasing distraction should disrupt counterargument production.
3. For junior-high students, effects should be greatest in the low distraction condition. These younger students should have less information about drug abuse and less developed attitudes than older teenagers and parents. Consequently, they should exhibit low levels of counterarguing and connections, and distraction should dampen messages' effects further.

## Research Procedures Used in Study One

The first large-scale experiment was conducted in Palo Alto, California (n=438). Subjects included three population groups: junior-high-school students (n=90), senior-high-school students (n=91), and parents (n=257). Subjects from each group were told that they were participating in a "Television Violence-Humor Project," and their opinions were sought in this area. Subjects viewed a 17½ minute videotape of a popular television program, which contained one of the test messages, as well as other (commercial) advertisements. There were three levels of distraction: none, low, and high. Low and high levels were differentiated by decibel levels of a tape recording during the experiment (a male-female conversation).

Following exposure, subjects completed post-exposure questionnaires. They were asked if they remembered any of the commercials, and if so, "what they thought about" while watching. The latter question provided data on cognitive responses--counterarguments and connections--to be reported here. Other questions concerned ranking of drug abuse as a social problem (providing a listing of 15 social problems), and interest in receiving booklets related to all of the commercials embedded in the programming. Two of the booklets were referred to in the anti-drug abuse commercials to which subjects had been exposed.

## Effects of Inductions

In persuasion research, level of distraction cannot be so dominant as to greatly reduce message comprehension. In the present research, effects of distraction were gauged through a self-report distraction scale, and through analyses of interactions of ad stimulus and recall with distraction level.

Results showed that the distraction induction was successful. The effect of distraction on recall varied markedly by ad stimulus and subject population, suggesting that distraction did not simply "wash out" opportunities for learning, but produced different results in interaction with other variables.

## Cognitive Responses: Counterarguing and Connections

Since small sample sizes preclude reliable analysis of three-way interactions between subject population, distraction, and ad stimulus, data are presented by subject population and ad stimulus (Table 1) and by subject population and distraction (Table 2). In both cases, for those respondents in each group who indicated that they recalled the ads, "percent recalling"



TABLE 1  
COGNITIVE RESPONSES, RANKING DRUG ABUSE AS A SOCIAL PROBLEM,  
AND BOOKLET REQUESTS, BY SUBJECT POPULATION AND AD STIMULUS

Cognitive Responses	Juniors				Seniors				Parents			
	WO*	BB*	P*	X	WO	BB	P	X	WO	BB	P	X
% commenting	72%	77%	85%	78%	90%	63%	68%	74%	92%	85%	78%	85%
% counterarguments	21	9	12	14	52	33	45	43	52	24	25	34
% connections	24	36	8	23	16	17	0	11	29	27	7	21
% other comments (n, base recall)	28 (29)	32 (22)	65 (26)	42	22 (31)	13 (24)	23 (22)	19	11 (79)	34 (71)	11 (79)	19
Ranking of Drug <sup>1</sup> Abuse as Social Problem	7.00	5.18	5.28		7.42	7.00	8.58		5.11	5.99	6.06	

(rg=1=15, most-least important)

Booklet Requests  
(higher mean = greater interest)

Booklet #1	n.s.				n.s.				n.s.			
Booklet #2 <sup>2</sup>	5.26	5.50	4.14		6.61	4.64	3.97		7.50	6.85	6.79	

1. (F = 2.62, 4 & 357 df, p < .05)
2. (F = 2.01, 4 & 357 df, p < .10)

\*WO = Walkout  
BB = Big Brother  
P = People

TABLE 2  
COGNITIVE RESPONSES AND RANKING DRUG ABUSE AS SOCIAL PROBLEM  
BY SUBJECT POPULATION AND DISTRACTION LEVEL

Cognitive Responses	Juniors				Seniors				Parents			
	none	low	high	x	none	low	high	x	none	low	high	x
% commenting	63%	67%	60%	63	62%	67%	59%	63	77%	78%	56%	70
% counterarguing	13	10	17	13	45	37	34	39	40	35	23	33
% connections	23	17	17	19	10	10	10	10	23	21	9	18
% other comments base recalling n	27 (30)	40 (30)	26 (30)	31	7 (29)	20 (30)	15 (29)	14	14 (86)	22 (85)	32 (85)	23
Ranking Drug Abuse <sup>1</sup> as Social Problem												
(rg=1-5, most-least important)	5.43	5.70	6.98		7.66	7.83	7.58		6.49	5.79	4.91	

1. (F = 2.66, 4 and 356 df, p < .05)

refers to the proposition who chose to respond to an open-ended question concerning their thoughts while viewing.

In Table 1 data are shown which are consistent with our expectations. Students chose to comment less about their thoughts while viewing than did parents, and the average percent commenting for senior high students is inflated due to the very high percent (90%) which commented on the "Walkout" commercial, compared with the other two ads. As was expected and supported by the

data, senior high students were more negatively pre-disposed to anti-drug abuse information, and therefore counterargued more and "connected" less than the other two groups. Yet the "Walkout" commercial stimulated equal percentages of counterarguments for seniors and parents (52% of comments), and of the three ad stimuli, it produced the most counterargument for junior-high-school students.

As also expected, parents exhibited more connections than the student groups, when compared to seniors, and

this is also generally true when compared to juniors. However, juniors exhibit the greatest percent of connections to the "Big Brother" commercial. Compared with other ad stimuli and with other respondent groups, there is no doubt because "Big Brother" appeals to juniors by portraying a junior-high-aged boy interacting with his older brother. Interestingly, while the ad was intended primarily to influence seniors, only 7% of their comments indicated connections with "Big Brother." Connections for this ad were higher among the other subject populations. Yet in any case, "Big Brother" stimulated the least counterargument among seniors, compared with the other two ad stimuli.

The ad stimuli also produced different effects for subject populations in stimulating concern for drug abuse as a social problem, and in stimulating interest in receiving more detailed information. As expected, from a list of 15 social problems, drug abuse was ranked as less of a problem by seniors than by juniors and parents. While "Walkout" was a powerful ad, judging by recall and comment levels, it was not very effective in stimulating drug abuse ranking for juniors and seniors; however, it was most effective among its intended audience--parents. "Walkout" also had most impact on parents' desires to receive a booklet containing information about drug abuse, relative to other ads, and relative to the other subject populations. Interestingly, while "Walkout" had powerful effects on cognitive responses among seniors, but little effect on ranking drug abuse as a social problem, the ad was most effective for this subject population in stimulating desire for a booklet. The latter results must be interpreted with caution, however, since the effects on booklet interest only approached significance ( $F=2.01$ , 4 and 357 df,  $p < .10$ ).

By examining effects of distraction (Table 2), it appears that percent commenting and cognitive responses generally decrease as distraction increases--note that effects are particularly strong in the "high" distraction condition. Adults are particularly sensitive to distraction, perhaps because their making cognitive responses require some effort, which is particularly difficult to justify when the information conflicts with existing attitudes. Consequently, for ad viewing conditions with high distraction, adults simply "tune out" much more than do junior- or senior-high students. However, the ranking of drug abuse as a social problem markedly increases with distraction among adults--yet the ranking decreases for junior-high students, and changes very little for seniors. It would seem that adults "compensated" for their diminishing cognitive responses by ranking drug abuse as a more serious social problem.

It was expected that increasing distraction would disrupt counterargument production among senior-high students, who were expected to be most negatively predisposed to the messages, to be most motivated to counterargue, and to have most information with which to counterargue. However, the results indicate counterarguing decreases only slightly for this audience group. Perhaps distraction cannot effectively overcome relatively strong orientations regarding drug abuse among senior-high students. For junior-high students, distraction has little effect--in fact, counterarguments slightly increase with distraction, contrary to prediction. It may be that the curiosity of these younger students with the topic of drug abuse is sufficient to overcome distraction.

## Second Study

### Research Procedures Used in Study Two

A second large-scale experiment was conducted in Bakersfield, California ( $n=753$ ). The objectives of this study included testing reliability of the Palo Alto results, and examining effects of new variables,<sup>2</sup> including measures of behavioral consequences of exposure to the ad stimuli. A booklet about drug abuse was mailed to all subjects participating in the Bakersfield study within a few days after their participation in the experiment. Then, in follow-up telephone interviews conducted 10 to 14 days later, a sub-sample of adults ( $n=242$ ) was asked if they recalled receipt of any booklet and if so was it helpful to them, did they read most or all of it, and did they have a drug discussion during the week?

Regarding reliability of the results, data in Table 3 compare the Palo Alto and Bakersfield studies in terms of nine dependent variables, for "Walkout" and "Big Brother." While absolute levels differ for the two studies, the patterns and directions of differences are consistent, except in two cases: percent of connections, and ranking of drug abuse as a social problem. For example, "Walkout" stimulated fewer connections (17%) in Bakersfield, compared with the respondent connections (29%) in the Palo Alto study. Nonetheless, the overall patterns indicate a reasonably high degree of reliability for the measures and procedures.

Effects of Competing Messages. The data show that competing messages affected gross recall, compared to the control condition of no competing messages (Table 4). When no competing messages were run, about three-quarters of the subjects recalled the three test ads. However, when public service or over-the-counter (OTC) drug ads were "competing," recall was heightened for the "Big Brother" and "Walkout" commercials, but not for the more neutral "People" commercial.

Examining the data more closely (Table 5), apparently the points recall for "Big Brother" and "Walkout" was affected by the presence of competing ads, although points recall for "Walkout" when OTC drug ads were present is slightly less than when no competing ads were present. However, points recall markedly increased when other public service ads were present.

Cognitive Responses: Counterarguing and Connections. Competing message conditions did not significantly affect counterarguing and connections. Moreover, competing message conditions did not produce differences in behavioral acts following exposure (e.g., requesting booklet, reading it, etc.). It appears, then, that "competing" messages simply increase recall, perhaps because the "competition" sensitizes people to the point made in the anti-drug abuse message.

<sup>2</sup>In the Bakersfield Study, a "competing messages" condition was implemented, in which subjects saw anti-drug abuse messages in the same program segments with over-the-counter drug ads, or other public service messages. The cover story, test procedures, audience populations, etc., were the same as in the Palo Alto experiment. However, based on the Palo Alto results, one level of distraction was used, midway in decibels between the "low" and "high" distraction levels in the earlier research.

TABLE 3

COMPARISON OF VARIOUS RESULTS FROM TWO EXPERIMENTS  
FOR TWO AD STIMULI

	<u>Palo Alto Study</u>		<u>Bakersfield Study</u>	
	<u>Walkout</u>	<u>Big Brother</u>	<u>Walkout</u>	<u>Big Brother</u>
<u>Recall</u>				
Points made	.41	.75	.25	.71
Situation	1.35	1.16	.84	.75
Sponsor	1.05	.91	.81	.79
<u>Cognitive Responses</u>				
% commenting	92%	85%	81%	80%
% counterarguing	52	24	52	15
% connections	29	27	17	26
Rank of Drug Problem (1-15)	5.12	5.98	5.23	4.82
<u>Booklet</u>				
Interest: #1	7.37	6.82	6.42	6.14
Booklet Interest: #2	7.50	6.85	7.07	6.65

TABLE 4

PERCENT RECALL OF TEST COMMERCIALS FOLLOWING EXPOSURE  
BY COMPETITION CONDITION

<u>Competing Message</u>	<u>Big Brother</u>			
	<u>Brother</u>	<u>Walkout</u>	<u>People</u>	<u>Pillbox</u>
None	72%	73%	74%	75%
OTC drugs	81	84	71	71
Public Service	94	88	66	77
<u>Total</u>	82%	82%	70%	74%

TABLE 5

INTERACTIONS: RECALL VARIABLES  
(scale: 0-3 low-high)

<u>Competition x Ad</u>	<u>Points Recall</u>		<u>Sponsor Recall</u>	
<u>None</u>				
Big Brother	.54		.74	
Walkout	.35		.80	
People	.24		.80	
Pillbox	.46		1.08	
<u>OTC Drugs</u>				
Big Brother	.79		.97	
Walkout	.30		.88	
People	.29		.78	
Pillbox	.35		.86	
<u>Public Service</u>				
Big Brother	1.10		1.02	
Walkout	.97		.87	
People	.28		.68	
Pillbox	.28		1.02	

(F = 3.03, 6 & 557 df, p .01) (F = 1.80, 6 & 557 df, p .10)

\*Means are shown for only those conditions producing F-ratios at the .10 level or better.

Data in Table 6 show various behavioral consequences of exposure to the various ad stimuli. All of the subsample (242 adults) had been mailed in a booklet (whether interest in receiving one had been indicated in the post-test interview or not). Yet, however exposure to the booklet actually affected behavior depended on the particular ad the adults had seen in the experiment, 10 to 14 days earlier. That is, the test showed that compared to the control condition (subjects who had not been exposed to any of the ad stimuli in the experiment), those who had seen "Walkout" were far more likely to recall receiving the booklet (91%), and they were more likely to report having had a discussion about drugs in the ensuing week. (Due to timing of the phone interview, the discussion may have been a direct consequence of experimental exposure, i.e., occurring in the time between participation in the experiment and receipt of the booklet by mail). On the other hand, subjects who saw "Big Brother" were least likely of all subjects to recall receiving a booklet, but they were most likely to report having read most or all of it, and to have found it "very helpful." Fewer of these subjects reported having a drug discussion compared with those who saw the "Walkout" commercial.

TABLE 6

BEHAVIOR IMPLICATIONS: RECALL RECEIVING BOOKLET,  
READING IT, AND DISCUSSING DRUGS, BY AD STIMULUS,  
10-14 DAYS AFTER EXPERIMENT

(Bakersfield Experiment)  
(n=242)

	<u>Ad Stimulus</u>			
	<u>None</u>	<u>Walkout</u>	<u>Big Brother</u>	<u>People</u>
Recall booklet receipt	n 68% (40)	n 91% (40)	n 61% (46)	n 70% (47)
Read most or all	37% (27)	51% (37)	54% (24)	34% (29)
Had drug discussion this week	24% (29)	53% (30)	32% (25)	27% (29)
Found very helpful	39% (23)	45% (29)	55% (18)	45% (20) (145)

Third Study

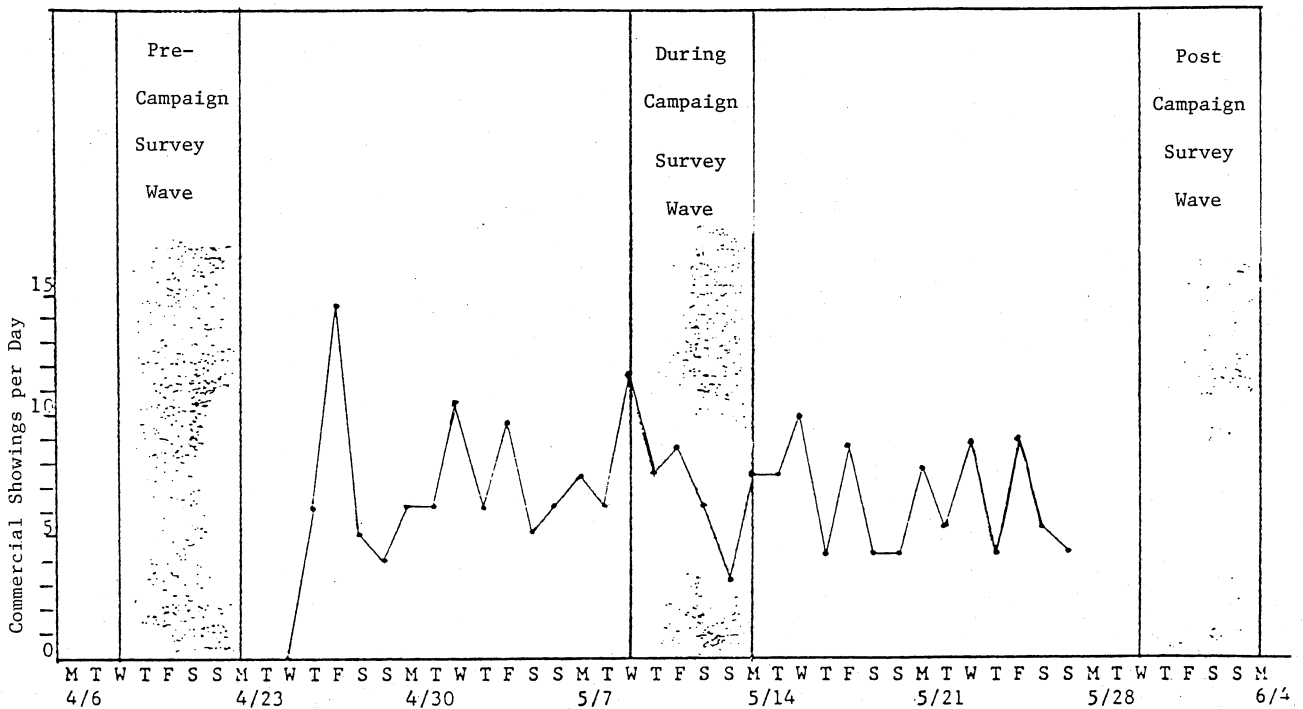
Research Procedures Used in Study Three

The final stage of the project involved a field experiment in which "Walkout" and "Big Brother" commercials were broadcast over a split-cable television facility in a west coast city.<sup>3</sup> In terms of the project objectives, the goal of this study was to test the response patterns obtained in the artificial laboratory environments to determine if these patterns would be validated in the natural conditions afforded by a field experi-

<sup>3</sup>The facility is run by AdTel, Inc. As illustrated in Figure 1, there were 209 simultaneous showings of the two commercials on Cable A and Cable B over a 32-day period. This averages 6.5 showings per day, equivalent to an \$18 million yearly advertising campaign. The cables reach highly comparable samples in terms of demographic characteristics.

FIGURE 1

SCHEDULING OF COMMERCIALS AND SURVEYS: FIELD EXPERIMENT



ment. Space considerations in this paper preclude full analyses of the procedures and results for the field experiment. Data in Table 7 essentially confirm the earlier laboratory analyses of relative differences between "Walkout" and "Big Brother" (see Studies One and Two) in terms of the effects of these two commercials on adults. Midway during the field test campaign, following essentially equal exposures to the two ads, "Walkout" was better recognized, stimulated recall of seeing any anti-drug ads (even if "Big Brother" were not recalled specifically), incurred more negative affect and counterarguing, and produced fewer connections than "Big Brother." In every case, compared with "Big Brother," "Walkout" was associated with relatively greater incidence of behavioral effects, e.g., reading about drug abuse, talking about the problem. These results were highly consistent with the Palo Alto and Bakersfield studies.<sup>4</sup>

Discussion

In this project report a summary has been presented of the rationale, procedures, and key findings of a major research project. We feel the study suggests some important areas of relevance for application in the public health area, as well as other managerial areas.

<sup>4</sup>

In the post-communication campaign wave of interviews, most effects were found to converge for the two ads, to pre-campaign levels, or lower. Intriguing questions for research remain concerning these post-campaign effects, resembling, in a general way, the well known "sleeper effect". The topic is essentially important when stimuli are repetitive, as in the case of advertising.

TABLE 7

FIELD EXPERIMENT RESULTS  
-During Campaign Wave Only-  
(Parents Audience)

	Walkout (Cable A)	Big Brother (Cable B)
$\bar{x}$ reported broadcast exposure	3.12	1.47
seen any anti-drug ads	66%	47%
best remembered	13%	11%
<u>Cognitive responses</u>	34%	62%
General positive affect	17	4
General negative affect	27	2
Specific counterargument	20	28
Connection		
	2.26	2.52
Rank of Drug Problem (1-17)	2.26	2.52
<u>Other Effects</u>		
In the past 2 days...		
$\bar{x}$ times thought about drug abuse	3.30	2.68
$\bar{x}$ times read about drug abuse	2.11	1.55
$\bar{x}$ times talked about drug abuse	2.55	2.01

First, the reliability, validity, and sensitivity of the technique described in the two laboratory studies are strengthened by the field experiment results. Existing pretesting facilities could readily adapt the procedures documented in this research.

We feel the data argue for the procedures illustrated here. While further research is needed to more pre-

cisely assess the effects of distraction and competing messages, these manipulations are representative of real world viewing, and they affect responses to advertising in important ways.

We also feel the lab results regarding cognitive responses during communication--counterarguments and connections--suggest that communication planners should be aware of these message effects. For example, a message could obtain a high day-after recall score in a test market, but people may recall the message because they counterargued against it.

We have not meant to imply that counterarguing is always "bad". In fact, depending on campaign objectives, counterarguing could productively "set people up" for messages run in the next campaign wave which effectively refutes these counterarguments.

Finally, we would argue for the kind of "selection scheduling" pretesting illustrated here, in which multiple cognitive and behavioral responses are measured. The kinds of responses should be beyond simply asking for people's attitudinal reactions to test ads. Such measures are highly reactive and, in any case, people cannot reliably assess an ad's impact on their future behavior.

Depending on campaign objectives, somewhat different dependent variables could be employed which are different from those used here. Furthermore, we feel the procedures and measures could readily be adapted for use with materials other than finished commercials--perhaps roughs or even storyboards. The procedures could also be adapted for use with advertising in other media.

#### REFERENCES

Robert S. Baron, Penny H. Brown, and Norman Miller, "The Relation Between Distraction and Persuasion," Psychological Bulletin, 30, 4(1973), 310-323.

Leon Festinger and Nathan Maccoby, "On Resistance to Persuasive Communications," Journal of Abnormal and Social Psychology, 68(1964), 359-366.

L. Johnston, Drugs and American Youth (Ann Arbor: Institute for Social Research, 1973).

Herbert C. Kelman, "Attitude Change as a Function of Response Restriction," Human Relations, 6(1953), 185-214.

Herbert E. Krugman, "Processes Underlying Exposure to Advertising," Proceedings of the Fourteenth Annual Conference of the Advertising Research Foundation, a(October 15, 1968), 14-19.

Herbert E. Krugman, "The Measurement of Advertising Involvement," Public Opinion Quarterly 32 b(1968), 583-596.

National Commission on Marijuana and Drug Abuse, Marijuana: A Signal of Misunderstanding (Washington: United States Government Printing Office, 1972).

Donald Roberts and Nathan Maccoby, "Information Processing and Persuasion: Counterarguing and Behavior," in Peter Clarke, (ed.), New Models for Communication Research (Beverly Hills: Sage Publishing Company, 1974), 269-302.

Peter L. Wright, "On the Direct Monitoring of Cognitive Responses to Advertising: A Necessary Reorientation in

Advertising Research," paper presented to Consumer Information-Processing Workshop, November 1972.

CONSUMER RESPONSE TO SEAT BELT USE CAMPAIGNS AND INDUCEMENTS:

IMPLICATIONS FOR PUBLIC HEALTH STRATEGIES

Leon S. Robertson  
Insurance Institute for Highway Safety

Abstract

Efforts to reduce motor vehicle deaths and injuries in the U.S. recently have concentrated on belt use advertising campaigns, buzzer-light reminder systems and starter-interlock belt systems in cars. Carefully controlled studies of advertising campaigns find no effect on belt use. The buzzer-light system had no effect and the interlock's effect was temporary and provoked adverse consumer reaction. Strategies requiring no action by the person to be protected have been available but neglected.

In the U.S. injuries in motor vehicles are the leading cause of death of persons aged 1-35 years. More than half of traumatic spinal cord lesions that result in paralysis occur in motor vehicle crashes (Kraus, et al., in press). About 3,927,000 motor vehicle related injuries were estimated from the National Health Survey in 1973 - almost 1 in every 50 U.S. residents (Health Resources Administration, 1974). Motor vehicle occupants sustain about three-fourths of fatal injuries with the remainder distributed among pedestrians, motorcyclists, and bicyclists.

The incidence and severity of vehicle occupant injuries would be reduced substantially if people would use seat belts and shoulder harnesses available in most cars and some other vehicles. However, attempts to increase belt use have had, at best, temporary and, usually, no success.

Media Campaigns

We all remember the "Buckle Up for Safety" and similar sloganeering campaigns in electronic and print media. Until recently these campaigns were launched with no precampaign experimentation to estimate potential effectiveness. Evaluation of effectiveness was limited to before and after claimed use in questionnaire surveys in a few cases. However, a 1969 study compared actually observed belt use with claimed use and found that claimed use exceeded actual use substantially (Waller and Barry, 1969). That type of error plus the lack of control groups in earlier studies led to the conclusion that the effect of media campaigns was unknown. A 1970 survey of drivers actually observed in their cars found only about 22 percent use of any belts in late model (1968-71) cars in metropolitan areas, less in older cars in metropolitan areas, and less than 10 percent use in small cities (Robertson, et al., 1972). It was evident that if earlier campaigns had any effects at all, a large potential for reduced injury through increased belt use remained.

Based on correlates of belt use in the 1970 study, six television messages were produced, each directed at a specific audience. One was judged best among 30 public service entries in the TV-Radio Advertisers Club of Philadelphia competition and another was among 10 finalists in 400 entries to the public service competition of the Advertising Club of New York. By arrangement with advertisers and a cable television company, messages were shown in prime time on one cable of a

split cable system designed for test marketing. The households on the cables are distributed on alternate streets in such a way that classical experimental-control design was achieved.

Each message was shown during programs appealing to audiences that were the targets of the messages (children, teenagers, fathers and mothers). The messages were shown 943 times over a period of nine months - the equivalent of a \$7 million campaign on a national basis.

Daily belt use observations were conducted on a rotation basis among sites throughout the test city for 11 months, including one month prior to and one month after the campaign. Belt use observations were matched to experimental and control cables by matching vehicle license numbers to motor vehicle administration files, obtaining owner names and addresses, and subsequently matching these names and addresses to those in cable company files. The observers did not know the purpose of the study and the people in the community did not know about the split cable system. Thus, a double-blind experiment was accomplished.

Comparison of belt use by drivers from households on the experimental and control cables, as well as with those on neither cable, revealed no effect on belt use of the television campaign (Robertson, et al., 1974). Similar results were obtained in a separate study using radio and television in some communities and none in others (Fleischer, 1972). Unless and until a rigorously designed study demonstrates otherwise, there is formidable evidence leading to the conclusion that mass media campaigns are ineffective means to increase belt use.

Mechanical Inducements

The federal agency responsible for reducing motor vehicle related injuries, the National Highway Traffic Safety Administration (NHTSA), attempted to require "passive" protection for vehicle occupants in 1972 and later model vehicles in frontal barrier crashes up to and including 30 miles per hour. "Passive" public health strategies refer to those approaches that do not require each individual to take action to protect himself (Haddon and Goddard, 1962). Passive protection is automatic - as in purified water, pasteurized milk, shielded electrical cables, fuses and circuit breakers in electrical systems, and inflatable air cushions in severe motor vehicle crashes.

Car manufacturers opposed the proposed federal requirement for passive vehicle occupant protection. The standard was then revised to allow a buzzer-light belt use reminder system as an alternative, optional to the manufacturers. With the exception of a few thousand cars equipped with inflatable air cushions, all cars manufactured for sale in the U.S. from January 1, 1972 through August 14, 1973 were equipped with the buzzer-light system. However, a study in the spring and

and summer of 1972 found no statistically significant difference in observed belt use between 1972 model year cars with and those without the buzzer-light system (Robertson and Haddon, 1974).

This study was also double blind. The observers did not know the purpose of the study and drivers did not know belt use was being observed. Cars with and without buzzer-light systems were identified by matching license numbers to vehicle identification numbers in motor vehicle administration files.

Two marketing studies of the buzzer-light system prior to its requirement claimed that it was effective in increasing belt use. However, one study was based on questionnaires from government employees using government owned cars in which belt use was mandatory by administrative directive (Perel and Ziegler, 1971). The second study involved a 30 day test drive by drivers who previously claimed not to use belts in a telephone survey. These drivers were given "intensive introduction to the vehicle, complete with test drive and a thorough explanation of the operation of all features which it contained" (Shaw, 1971). After 30 days, an interviewer took a drive with the driver and observed belt use. Apparently the researchers unknowingly rediscovered the Hawthorne effect.

Cars manufactured for sale in the U.S. after August 14, 1974 were required - if the manufacturer chose not to install the specified passive protection - to install a system that would not allow the car to start if front seats were occupied by more than a certain weight and belts were not extended or latched. Again, with the exception of a few thousand air cushion equipped luxury models, the so-called interlock belt use system was installed in 1974 model year cars.

Initially belt use was found to be nearly 60 percent in interlock equipped vehicles (Robertson, in press). A survey of owners of 1974 cars found only 29 percent mentioned the interlock among the three features liked least about their new cars (Robertson, 1975a). However, many of this minority objected strongly (Congressional Record, 1974). Letters to congressmen resulted in the enactment of a federal law, subsequently signed by the President, banning the interlock as well as continuous buzzer-light systems.

In the spring of 1975 a survey involving observed belt use in four metropolitan areas found that two-thirds of drivers in 1974 and 1975 model cars were not using any belts and only 27 percent were using shoulder harnesses in addition to lap belts (Robertson, 1975b). Because of lower use in small cities and towns, lower use among children, and lower use in cars as they get older, the use of belts over the lifetime of these cars will be less than 20 percent.

#### Mandatory Belt Use Laws

Another means proposed to increase belt use in the U.S. is to enact laws requiring use. A few countries now have such laws following their reported effectiveness in Australia. Belt use increased to more than 60 percent in rural areas and 70 percent in urban areas in the first year following enactment of a mandatory use law in the state of Victoria. This increase was accompanied by a 21 percent reduction in vehicle occupant fatalities in metropolitan areas and a 10 percent reduction in rural areas (Foldvary and Lane, 1974). Whether the initial belt use achieved can be maintained over the long run remains to be seen.

The only U.S. jurisdiction thus far that enacted a belt use law, Puerto Rico, had only a small, temporary

increase in use. After reaching a maximum of 24 percent, public reaction resulted in a reduction in the penalty for conviction and belt use fell to 10 percent (Insurance Institute for Highway Safety, 1975). Despite endorsement by the major auto manufacturers, belt manufacturers, NHTSA, state and national safety councils, and others, no state in the U.S. has enacted a mandatory belt use law. Enactment by a large number of states, much less enforcement of and compliance with such a law, remains problematic.

#### Public Health Strategies

The history of public health suggests that attempts to control damage to human beings by changing everyone's behavior is never completely successful. Even when the behavior is required by law or administrative directive, such as immunizations to enter school, compliance seldom exceeds 80 percent. Campaigns to increase examination for problems such as cervical cancer seldom achieve more than 50 percent success (Robertson and Heagarty, 1975). And, as I have noted, changing behaviors that involve frequent attention and inconvenience such as belt use, is especially difficult.

With present technology, behavior change strategies may be the only ones available for some public health problems. However, by focusing on behavior change, we have sometimes delayed the adoption of available alternative strategies or neglected their development. In the case of motor vehicle occupant crash protection, consumers have been needlessly badgered by slogans, buzzers, lights, and interlocks while alternative passive approaches were left on the shelf.

The air cushion that inflates automatically in a severe crash, gently controlling the deceleration of occupants, is but one of a number of such approaches. The air cushion system has been installed in more than 8,000 cars that have now accumulated well over 100 million miles of on-the-road travel. Hundreds of laboratory tests have been done and more advanced systems than those now on the road are available. Only two inadvertent deployments have occurred while the vehicles were being driven, causing only minor injury. More than 1,000 minor crashes have occurred in which the air cushion did not deploy. In severe crashes where the vehicle was towed away, death or life threatening injury occurred in less than 3 percent of occupants of air cushion equipped vehicles compared to 5 percent using belts in belt equipped cars and 11 percent not using belts in similar cars in similar crashes (Robertson, 1975b).

Yet how many consumers know about the availability of these air cushion systems? Only token advertising marked their introduction with little or no followup. A film demonstrating their effectiveness and convenience was shown to dealers but was not even made available to the public.

Other passive approaches have been neglected as well. Our roadsides are lined with trees, poles and other solidly built structures that should be removed or modified to yield gently when they are hit by vehicles that stray only a few feet from the roadway. At the same time most of our vehicles are unnecessarily capable of speeds twice the national 55 mile an hour speed limit at a waste of scarce fuel and raw materials as well as lives.

Does it make sense to continue mainly to attempt to change every individual's behavior whether by advertising, education, or law enforcement? Or does it make more sense to force decision makers to provide automatic protection where the technology is available to

do so? If faced with the choices today, would we purify water at the source or would we launch an ad campaign to attempt to persuade everyone to boil his or her drinking water? Would we require pasteurization of milk before it is sold or would we pass a law that each family had to boil their milk before it is consumed? The choices we face today with respect to motor vehicle injuries, as well as other public health problems, are no different than these.

#### References

Congressional Record, "Motor vehicle and schoolbus safety amendments of 1974," (August 12, 1974), H8119ff.

G. A. Fleischer, An Experiment in the Use of Broadcast Media in Highway Safety (Los Angeles: Department of Industrial and Systems Engineering, University of Southern California, 1972).

L. A. Foldvary and J. C. Lane, "The Effectiveness of Compulsory Wearing of Seat-Belts in Casualty Reduction," Accident Analysis and Prevention, 6 (September, 1974), 59-81.

W. Haddon, Jr. and J. L. Goddard, "An Analysis of Highway Safety Strategies," In Passenger Car Design and Highway Safety (New York: Association for the Aid of Crippled Children and Consumers Union of the U.S., 1962).

Health Resources Administration, Current Estimates from the National Health Survey, United States - 1973 (Washington, D.C.: U.S. Government Printing Office, 1974).

Insurance Institute for Highway Safety, "Belt Use in Puerto Rico," Status Report, 10 (May 12, 1975), 4.

J. F. Kraus, et al., "Incidence of Traumatic Spinal Cord Lesions," Journal of Chronic Diseases, in press.

M. Perel and P. M. Ziegler, An Evaluation of a Safety Belt Interlock System (Washington, D.C.: National Highway Traffic Safety Administration, 1971).

L. S. Robertson, "Safety Belt Use in Automobiles with Starter-Interlock and Buzzer-Light Reminder Systems," American Journal of Public Health, in press.

L. S. Robertson, "Factors Associated with Safety Belt Use in 1974 Starter-Interlock Equipped Cars," Journal of Health and Social Behavior, 16 (June, 1975a) 173-177.

L. S. Robertson, "Motor Vehicle Occupant Restraint Use and Effectiveness in Real-World Crashes," Appendix A in testimony of W. Haddon, Jr. and A. B. Kelley before the National Highway Traffic Safety Administration, Public Meeting on Occupant Crash Protection, May 19, 1975b.

L. S. Robertson, et al., "A Controlled Study of the Effect of Television Messages on Safety Belt Use," American Journal of Public Health, 64 (November, 1974), 1071-1080.

L. S. Robertson and W. Haddon, Jr., "The Buzzer-Light Reminder System and Safety Belt Use," American Journal of Public Health, 64 (August, 1974), 814-815.

L. S. Robertson and M. C. Heagarty, Medical Sociology: A General Systems Approach (Chicago: Nelson-Hall Company, 1975).

L. S. Robertson, B. O'Neill, and C. W. Wixom, "Factors Associated with Observed Safety Belt Use," Journal of Health and Social Behavior, 13 (March, 1972), 18-24.

D. J. Shaw, "Interim Results from Test Drive I Advanced Features Study." Transmitted to Docket 69-7, National Highway Traffic Safety Administration, July 27, 1971.

P. F. Waller and P. Z. Barry, Seat Belts: A Comparison of Observed and Reported Use. (Chapel Hill: University of North Carolina Highway Safety Research Center, 1969).



## COMMUNICATION ISSUES IN DIFFERENT PUBLIC HEALTH AREAS

F. Gerald Kline, University of Michigan  
Peter V. Miller, University of Michigan  
Andrew J. Morrison, Market Opinion Research

### Abstract

This paper deals with information selection by adolescents in two cities from several media concerning three public health topics: drug use, alcohol use, and family planning. A three factor ANOVA design with repeated measures on two factors is used to investigate the adolescents' message discrimination on the topics. We find significant topic, channel, and city differences in this behavior, as well as city-topic and topic-channel interactions. Implications of these findings for policy makers are discussed.

### Introduction

The interest in "information processing" aspects of consumer activity in recent years has paralleled a similar development in the study of mass communication. Dissatisfaction with the traditional view of the media's workings in society -- the so-called "hypodermic needle" metaphor -- has led to a more intensive consideration of what uses audience members make of media offerings, rather than on the characteristics of communicators or messages which might make them powerful persuaders. Both theoretical reasoning and empirical productivity impel this recent orientation. The former treatment of the audience as an homogeneous, passive collection of individuals left much to be desired in the way of explanation for their perverse tendency to ignore, or misconstrue, varied communication efforts (Bauer, 1964). When the null hypothesis on media effects continued to prevail in areas like political persuasion research, furthermore, it became apparent that new tactics -- including a recognition of audience potency in the communication process -- were required to search out the effects that did exist. Hence, we find more attention being paid these days in the mass communication research literature to the potential receivers of mass media messages, focussing on their interpersonal environments and intrapersonal predilections and abilities.

One salient on this front has been a renewed effort to identify or conceptualize the factors of interest in the communication process. Some of this work has concentrated on criterion variables, as in the recent research on "agenda-setting" by the media in the political arena. Here new dependent variables stress the media's potential ability to structure the ground rules of debate rather than the media's supposed persuasive power. [The media may not be able to tell us what to think, but they may tell us what to think about (Cohen, 1963).]

Agenda-setting research forces us to consider, from the audience's point of view, those rather broad pieces of information which may be culled from the media concerning societal problems. It is the audience member who elucidates the media's potential effects by the way he cuts up the big issues of the day. This is in sharp contrast to the research of earlier days which focussed on the communicator or aspects of his message to define the universe of possible effects.

Another audience-based approach has been pioneered in England and Israel. The "uses and gratifications" perspective on mass communication examines the reasons why people choose with media content and how they use that

content (Blumler and Katz, 1974). The typical flow of causality from media stimulus to audience response is simply reversed, in this case, to emphasize the receiver's needs as the force behind his choice of certain communicator offerings. The approach also gives us a better handle on possible media effects, in that the audience predispositions may be considered as mediating variables which soften or enhance the import of a particular communication (Kline, Miller and Morrison, 1974).

In this paper, we consider yet another reconceptualization which has followed from the emphasis on the receiver in the communication process. Message discrimination, as a measure of media exposure, incorporates different receiver processes than the traditional "time-spent" estimates. It also fits nicely into the information processing orientation of many consumer researchers by highlighting one of the stages in that process -- the input or "first order" stage of selecting pieces of information for further consideration (Ray and Ward, 1975).

Message discrimination tells us something about both the receiver and the sources and subjects of his reception. We learn about audience priorities for subject-matter and channels of reception, as well as the frequency and strength of messages in various channels. Disentangling the contributions of the receiver's priorities and the availability of information on certain topics in certain channels is not a simple task, and one we will not completely resolve in this paper. Rather, our analysis leads us to speculate on the channel, topic, and "media environmental" differences which may, in conjunction with receiver preferences, account for variance in the way teenagers select information from the mass media on several public health topics.

The subject-headings of interest are drug use, alcohol use and family planning. We have included an additional topic area in the analysis for comparison purposes, occupational information. In examining the amount of information selected from various media on these topics by adolescents, we hope to elucidate the interface between the macro-perspective of the communicator or public health policy maker and the micro-orientation of the teenage audience member in the transmission of public health information.

### Study and Analysis Design

The data presented in this paper come from two waves of interviews undertaken as part of a year-long study of the social context of adolescent media use. Nearly three hundred households containing at least one adolescent between 14 and 17 years of age were samples in each of two cities (Flint, Michigan and Toledo, Ohio) which we attempted to match on media market characteristics. Three random groups were formed in each city so that we could have control and experimental groups and additional controls for pretest sensitization in a multi-wave field experiment. The resulting scheme, similar to the Solomon Four-Group design (Campbell and Stanley, 1966), was used to assess the impact of messages about family

planning, alcohol use, drug use and occupational information. Teen radio stations were the vehicles for these experimental messages, which we produced.

In this report, we will restrict ourselves to non-experimental aspects of the study and concentrate on making topic, city, and channel comparisons of message perception and recall by the adolescent samples. Reports on other aspects of the larger project may be found in Kline, Miller and Morrison (1974), Miller, Morrison, and Kline (1974), Miller, Kline and Morrison (1975), and Morrison, Kline and Miller (1975).

Our analysis focusses on city, channel and topic differences in message discrimination. We chose a three factor ANOVA design with repeated measures on two factors to investigate these effects. The choice of design was indicated by the fact that our sample of respondents had been interviewed about all four topics, and had reported on their media message discrimination across seven channels within each topic. Thus, we treated the topic-channel survey questions as repeated measures on the topic and channel factors. The third factor, city of residence, naturally did not contain a repeated measures component. (See Winer, 1962, pp. 319-337 for details of the analysis procedure).

### Concepts and Hypotheses

#### Message Discrimination

This concept was introduced into the literature of mass communication research by Clarke and Kline (1974) as a replacement for the standard "time-spent" measures of media exposure. The measure entails asking a respondent in a survey interview what he has seen or heard about a particular topic in a particular medium within a particular time frame (the maximum being one month). The interviewer records the reply and probes to see if any additional thoughts can be obtained, and the process is repeated for each of the media under consideration. The open-ended responses are coded into "messages" according to an inductively-derived coding scheme, which allows for several messages to be coded per question. An overall message discrimination score is constructed for each respondent by summing all of the messages coded for all topics and media.

Thus, we conceive of message discrimination as the selection of units of content for attention and recall out of a universe of competing stimuli. In order to evidence media exposure, the respondent must recall something he actually saw or heard while in contact with the media. The emphasis is on actual, or engaged exposure, rather than potential exposure addressed in measures of how much time was spent with a given medium during a given period. The discriminations are considered first-order cognitive processing, a necessary but not sufficient condition for integration of information into a larger cognitive structure (Kline and Davis, 1973).

#### Topics

The topics chosen for the study reflected our judgment about subject-matter which would be salient for a teenage audience, and concentrated on public policy issues whose unique character might be investigated. The three public health areas, for example, have been highlighted as major problems among youth in American society. Drug use is a familiar and seemingly ubiquitous issue, while alcohol use has recently become a source of even more concern to some public health experts. The shortage of correct information among adolescents regarding family planning practices is well documented, (Weiss, et al., 1974) and the taboo character of this subject-matter adds another interesting dimension from a public policy perspective. Rather than consider these three topics in isolation,

we chose to offer a comparison with occupational information which, while perhaps equally important to teenagers, does not have the "health problem" character.

#### Channels and Cities

The media we investigated included a range from the truly "mass" type to more specialized channels; the list includes television, newspapers, magazines, radio, books and pamphlets, film and billboards. Our anticipation was that this variation in media would allow us to make some inferences about the informativeness of the channels across the different topics. Since the group contains some media usually not considered in the context of public health communication research, we may also shed some light on the nature of these channels in this specialized area. The cities were chosen by us both for their geographical location for research purposes, and because of the kinds of media mix which existed in them. We hoped for a relatively good matching, such that the two places would be approximately equal in the amount of information in their environments about the topics we considered. Furthermore, we wanted to find places which would not cross-feed information to one another. While we were successful to a degree in these goals, it remained important to consider the possibility of city differences in attention to the topics (differences in media environment") or differences in the populations studied which might produce differential salience for individual topics. From a policy view point, city differences would indicate a necessity to consider these factors when planning public information campaigns and research on them.

#### Hypotheses

We would expect the combination of audience priorities and channel and topic penetration to result in differences in message discrimination by these two factors. It is likely that teenagers pay greater attention to certain channels (e.g., television) and have greater interest in certain topics (e.g., drug use). It is also likely that certain topics receive more "play" across channels, and that the frequency of information on all four topics would vary by channel.

Furthermore, we might anticipate topic by channel interactions, as some topic information is likely to be played, and received, in some channels more than others. To the extent that city differences occur, (we expect no differences) we would expect them to be the result of greater media attention to a particular topic in one city versus the other. City by topic interactions might also occur in such an event, but city-channel interactions would be quite unexpected.

#### Analysis Results

The results of the repeated measures analysis of variance are displayed in Table I. We find significant main effects for all three predictor variables -- city, topic, and media channels. Significant interactions occur between city and topic, and between topic and media channel-variables. We now turn to a more detailed look at each of these significant effects.

#### Cities and Topics

In designing a field experiment such as the one described here, we attempted to match two cities in terms of both existent mass media and demographic factors. We hoped, then, that the baseline message discrimination levels between the two cities for any particular topic would be minimally and non-significantly different. However, as shown in Table II, there is a significant overall difference in mean messages discriminated between the two cities as well as for overall topics except family planning. We

might interpret this as resulting from a higher message display in the Toledo media for alcohol, drug, and job topics. The small difference for the family planning topic could possibly be attributed to intensive efforts by the Mott Foundation family planning program to get family planning information disseminated in the Flint area.

Table I

Results of Three Factors Repeated Measures ANOVA Analysis

Source	df	MS	F
<u>Between Subjects</u>	155		
A (City)	1	14.0	6.10*
Subjects within Groups	154	2.26	
<u>Within Subjects</u>	4212		
B (Topics)	3	11.5	29.48**
AB	3	1.7	4.36**
B x Subj. w. Groups	462	.39	
C (Channels)	6	13.5	30.68**
AC	6	.75	1.70
C x Subj. w. Groups	924	.44	
BC	18	2.50	6.94**
ABC	18	.60	1.70
BC x Subj. w. Groups	2772	.36	

\* Significant at .05 level.  
 \*\*Significant at .01 level.

Topic differences are of more interest to us. It appears that occupational and drug messages are discriminated by adolescents much more than alcohol and family planning messages. Of course, this finding leads to a consideration of what topic characteristics might enhance information receptivity in the adolescent population. Certainly, sheer display of information might be a factor. Many kinds of messages, including want ads, television dramatizations and news stories may be perceived as occupational information, and so its "play" in the media is likely to be greater because of the size of the topic category.

On the other extreme, family planning information apparently still has something of a taboo character about it, especially as family planning matters are discussed (or not discussed) in the mass media. There appears to be more messages discriminated about drug use than the alcohol topic. This difference may be due to the greater news display about drug-related incidents (crimes, latest scientific findings, and the like) than one finds for alcohol related happenings. Moreover, messages about alcohol use may predominantly be tied to advertisements for alcoholic beverages or public service announcements warning against the use of those same beverages.

As mentioned earlier in the paper, it is difficult to disentangle the contributions of media emphasis on certain kinds of content and audience information priorities in producing a message discrimination score. Some control on content in the media would help, and where we have conducted content analyses of the channels (with regard to family planning), we have found a strong monotonicity between proportions of messages discriminated about the topic and media "play" during that time. Whether that finding would hold up for the other topics is unclear, since family planning is such a special kind of subject matter, and since our content analytic efforts have not yet extended to the other topic areas. Ideally, an experimental method would be employed in which content available could be strictly controlled

and measured, and audience priorities for information manipulated. The "signal detection" literature would provide a model for such an analysis (Swets, 1964). Meanwhile, in our "actuarial" accounting of message discrimination differences by topic, we are forced to concede the ability of both audience predispositions and sender emphases in flavoring the outcome.

Table II

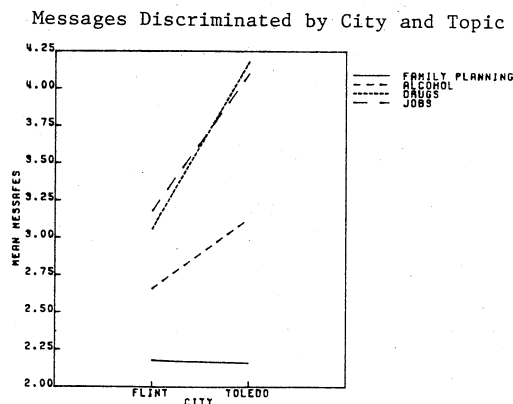
Mean Number of Messages Discriminated by Two City Samples Across Four Topics<sup>1</sup>

	Flint	Toledo	Overall Mean
Occupational Information	3.18 (N=78)	4.10 (N=82)	3.65 (N=160)
Drug Use	3.06	4.18	3.62
Alcohol Use	2.66	3.13	2.90
Family Planning	2.18	2.16	2.17
Overall Mean	11.07	13.58	

<sup>1</sup>City Effect significant at .05 level.  
 Topic Effect significant at .01 level.  
 City by Topic Interaction significant at .01 level.

The city by topic interaction is due, again, to the equal numbers of messages discriminated about family planning in the two cities as compared to the other three topics. This interaction is seen clearly in Figure 1. The key finding here is the topic differences that exemplify the necessity for message producers to view these topics differently in terms of designing mass media campaign strategies in different locations.

Figure 1



#### Media Channels and Topics

We have noted above some reasons why the clear differences in messages discriminated among the four topics may have occurred. It is now of interest to determine why, across topics, significant differences in mean number of messages discriminated occur among media channels. If one arrays the media from most to least total messages discriminated across topics, the array takes on a "mass -- 'focused' -- mass" media character.

The data in Table III show that television and newspaper dominate as primary message discrimination channels while two other mass media, radio and magazines, appear to have had the fewest messages discriminated in them.

Among these four mass media, this may reflect the news and, to a lesser extent public service characteristics of television and newspapers and the entertainment focus of many magazines and radio stations attended to by this adolescent sample.

Table III

Mean Number of Messages Discriminated in Four Topics Across Seven Media Channels<sup>1</sup>

	Occupational Info.	Alcohol Use	Drug Use	Family Planning	Total
Television (N=160)	.84	.81	.70	.60	2.95
Newspapers	.38	.46	.76	.39	1.99
Magazines	.36	.23	.28	.31	1.19
Radio	.34	.33	.45	.22	1.36
Books/Pamphlets	.63	.23	.58	.24	1.69
Films	.41	.44	.55	.14	1.54
Billboards	.68	.38	.32	.27	1.65
Total	3.65	2.90	3.62	2.17	

<sup>1</sup>Channel Effect significant at .01 level.  
Topic Effect significant at .01 level.  
Channel by Topic Interaction significant at .01 level.

The "focused" media of books and pamphlets, billboards, and non-television films falls in between these two extremes. We perceive these media to be "focused" relative to these topics because of single topic, narrow-theme nature of pamphlets, billboards, and school movies, in particular. Some of the message discrimination differences for both the "mass" channels and the "focused" media, naturally, may be due to some purposive use of channels on the part of adolescents for certain kinds of information. While serendipity may explain the sighting of a drug use message on television, an occupational information message from the newspaper is likely to reflect a scanning of the want-ads. Similarly, purposive use of the "focused" channels is likely for certain information, particularly in the case of books and pamphlets. We would expect adolescents to have subjective probabilities, in other words, on the likelihood of encountering various kinds of content in these channels, which may be reflected in the differences in message discrimination across channels.

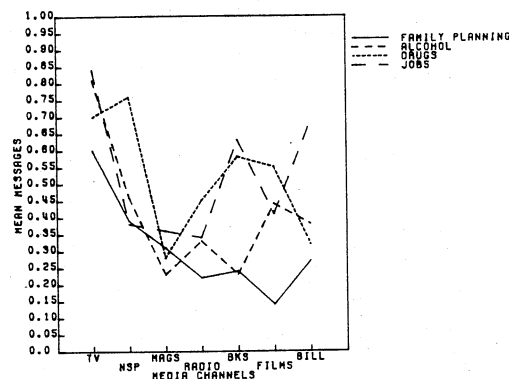
The topic-channel interaction is the joint product of message producer's current use of mass or focused media to display messages about particular topics and the media use habits of the adolescent audience. Occupational messages were picked up primarily from television, books and pamphlets, and billboards. On the first of these, of course, a great deal of occupational information is evident in the form of actors playing out a variety of occupational roles. Job messages seen on billboards can be attributed in the main to military advertising (Miller, Kline, Morrison, 1975), while occupational pamphlets are a prominent feature of most school counselor offices. The lack of messages seen in newspapers may indicate that 14-17 year old adolescents are not yet at a stage where they are very concerned about future job possibilities, or employment statistics. Drug topic

messages were most perceived in newspaper and television media and, to a lesser extent, in books and pamphlets and films. On the one hand, this may reflect wider news coverage of drug related items in these mass media news shows. On the other hand, we may also see the influence of school pamphlet and film campaigns on drugs, as well as drug use as a central theme in movies seen at local theaters.

Alcohol messages are seen extensively on television compared to all other media, perhaps as a consequence of public service announcements. This may hold true for the relatively few family planning messages discriminated. From the adolescent information user's viewpoint, parents and peers may be more likely choices as information sources about these topics. Figure II illustrates the channel differences in message discrimination and the channel-topic interaction.

Figure 2

Messages Discriminated by Topic by Media



Conclusion

It is clear that message discrimination is not the only criterion we would like to examine with regard to communication in the public health area. We might be interested, for example, in some 'higher-order' product of information processing, such as cognitive restructuring, new knowledge, or actual behavior shifts. The latter would be particularly of interest to policy makers, one expects.

At the same time, it is evident that a report of the actual information selection by teenagers concerning some pressing public health issues -- along with a consideration of some major factors which may account for a variance in this behavior -- has some merit in its own right. The initial selection, storage and recall of bits of information on these topics from the whelter of stimuli available in the media is a process which may, in large measure, define what higher order effects are possible. Without initial contact with the stimuli and some attention to them we cannot attribute to the media any shifts in knowledge or behavior which may occur in the population of interest. And we have seen how media exposure measures often used in mass communication research do not really capture the initial processing character of attention to and perception of certain media messages, but focus instead on "potential" exposure -- a self-report of time spent with a medium. Hence, it seems reasonable to examine message discrimination in the light of some geographic, subject-matter, and channel differences, and the interactions among them. In so doing, we bring both communicator policy decisions and audience selection priorities together in a single measure, and we make inferences about the macro- and micro- factors which determine initial information processing about these important topics.

## References

Raymond Bauer, "The Obstinate Audience: The Influence Process from the Point of View of Social Communication," American Psychologist 19(May, 1964), 319-328.

Jay Blumler and Elihu Katz, The Uses of Mass Communications: Current Perspectives in Gratifications Research (Beverly Hills: Sage Publications, 1974).

Donald T. Campbell and Julian C. Stanley, Experimental and Quasi-Experimental Designs for Research (Chicago: Rand-McNally, 1963).

Peter Clarke and F. Gerald Kline, "Media Effects Reconsidered: Some New Strategies for Communication Research," Communication Research, 1(April, 1974), 224-240.

Bernard Cohen, The Press, the Public and Foreign Policy (Princeton: Princeton University Press, 1963).

F. Gerald Kline and Dennis Davis, "Cognitive Similarity and Facilitation of Communication," Paper presented to The World Congress on Human Communication, Barcelona, Spain, November, 1973.

F. Gerald Kline, Peter V. Miller and Andrew J. Morrison, "Adolescents and Family Planning Information: An Exploration of Audience Needs and Media Effects," in Jay Blumler and Elihu Katz, eds., The Uses of Mass Communications: New Perspectives on Gratifications Research (Beverly Hills: Sage Publications, 1974).

Peter V. Miller, F. Gerald Kline, and Andrew J. Morrison, "Adolescents' Learning About Military Occupations in the Mass Media," Proceedings, Research Conference on the Social Psychology of Military Service, Chicago, April, 1975.

Peter V. Miller, Andrew J. Morrison and F. Gerald Kline, "Approaches to Characterizing Information Environments," Paper presented to the Annual Meeting, International Communication Association, New Orleans, April, 1974.

Andrew J. Morrison, F. Gerald Kline, and Peter V. Miller, "Aspects of Adolescent Information Acquisition on Drugs and Alcohol Topics," in Ronald Ostman and Hamid Mowlana, eds., Communication Research and Drug Education (Beverly Hills: Sage Publications, 1975).

Michael Ray and Scott Ward, "The Relevance of Consumer Information Processing Studies to Communication Research, 2(July, 1975), 195-202.

E. Swets, Signal Detection and Recognition by Human Observers: Contemporary Readings (New York: John Wiley, 1964).

Eugene Weiss, F. Gerald Kline, Everett M. Rogers, Mitchell E. Cohen and Jennifer A. Dodge, "Communication and Population Socialization among Youth: Communication Behavior and Family Planning Information," Paper presented to the Population Association of America, New York, April 1974.

B. J. Winer, Statistical Principles in Experimental Design (New York: McGraw-Hill, 1962).

# THE BEHAVIOR OF THE HEALTH CARE CONSUMER: A SELECTIVE REVIEW

Lawrence H. Wortzel, Boston University

## Abstract

Consumer behavior analysis and findings are an important input to the design of health care marketing programs. This paper is an attempt to present a framework for understanding consumer health care behavior, and to present selected findings. The paper concentrates on primary demand aspects and focusses on three types of physician visits: preventive, diagnostic, and therapeutic. A model is presented to predict behavior for preventive and diagnostic situations, and behavior in therapeutic situations is described and analyzed.

## Introduction

For a variety of reasons ranging from the purely humanitarian to the purely economic, marketing has become an accepted activity in many health care institutions and settings. The functions that marketing is expected to serve in the health care field are not dissimilar to the functions marketing is expected to fulfill in the commercial sector of the economy. And it is not surprising that the marketing problems faced by many health care institutions are similar to those faced by firms in the commercial sector. Perhaps some examples will make this point more clearly.

Beale and Schroeder (1973) describe the marketing efforts undertaken by a new urban health center. The marketing task was simply to register potential users with the center. An adequate level of registration was required in order to justify a full staff for the center. Registrants did not have to pay either a registration or a maintenance fee nor did they take on any other contractual obligation; registration merely indicated intent to use the center's services. The center was promoted primarily through direct mail advertising.

A Government technical assistance publication (U.S. Department of Health, Education, and Welfare, 1973) describes the marketing tasks involved in HMO marketing in similar terms, referring to media advertising and to personal selling efforts directed both toward organizations and toward individual potential subscribers within those organizations. Kotler (1975) lists health institutions including hospitals, HMO's and communities that could undertake health care marketing efforts. He goes on to describe the variety of marketing efforts that these organizations and institutions could undertake, including attracting patients, attracting donations of blood or money, attracting physicians, attracting contributions to medical causes, and encouraging behavior likely to be productive of better health. Invariably, these examples discuss the use of the usual marketing tools.

Virtually all marketers working in the commercial sector of the economy recognize that the basis of a marketing program is an understanding of the relevant aspects of consumer behavior. The emergence of a consumer behavior literature, consisting of both conceptual work and empirical findings organized and reported so that it is both accessible and useful to commercial marketers is a testimonial to this recognition. Moreover, the way in which this body of literature has developed has facilitated the conduct of additional research because of the relative ease with which a researcher can identify gaps in knowledge. Health care marketers also recognize

the importance of understanding consumer behavior. There is a growing body of literature relating to health care that can certainly be labelled "consumer behavior" literature. As yet, however, this literature has not been organized so that its concepts and findings can be easily used by health care marketers in building their marketing programs. And, possibly, the content and dimensions of this literature may not be well enough known among potential researchers in health care consumer behavior. This paper, therefore, will attempt to do two jobs: (1) to review some of the content and dimensions of the literature relating to health care consumer behavior, and (2) to organize this literature, and to develop a focus for it that will make somewhat easier the tasks of applying it to marketing problems and of identifying potentially worthwhile research projects. The review to follow is a very selective one; no attempt has been made to cover all of the literature. Its major purpose is to give interested health care marketers and researchers some guides and some entry points for a possibly more intensive search. The attempt at organization is an effort to cast the literature into a simple straight forward, marketing oriented and managerially useful framework.

## A Framework for Health Care Consumer Behavior

Exhibit 1 presents schematically the framework that will be used in organizing this paper. The words and concepts are familiar enough and appear, at least in part, in either oral or written publications; is by no means wholly original. Essentially, the framework is a 3 x 3 matrix in which the columns are types of medical care visits and the rows are consumer decisions. Thus, consumer can make a visit for preventive reasons (a polio shot), or for diagnosis (a "pap" smear), or for treatment (in response, say, to a pain in the neck). Similarly, the consumer has both a primary and a selective demand choice to make, both of which may be affected by the consumer's previous experience with visit types and with providers.

Each cell of the matrix encloses the factors that determine behavior with respect to each combination. These factors can be classified into four broad groups: people specific (such as ethnic background, social class, age, sex, life style); promotion specific (whatever an institution does to obtain and satisfy patients) "product service" specific (whatever it is that the consumer perceives is being provided); and institution specific (such as type of provider and location of the provider in relation to the consumer).

Since this paper had to conform to limitation in space, it has been necessary to concentrate the review on only certain cells of the matrix, and on only selected factors within each cell. In order to conform most closely to the theme of the session in which this paper will appear, the paper will concentrate on primary demand and on people specific factors.

## The Decision to Seek Care: Preventive and Diagnostic Visits

Kasl and Cobb (1965) provide rather good descriptions of preventive and diagnostic behavior, defining each

Exhibit 1

Health Care Consumer Decisions

Consumer Decision	Type of Visit		
	Preventive	Diagnostic	Therapeutic
Whether to seek care (Primary Demand)	*		
What Provider to choose (Selective Demand)			
Whether to return -- at all -- to the same provider -- to a different provider (post purchase satisfaction)			

- \* Determining factors:
- people specific
  - promotion specific
  - "service product" specific
  - Institution specific

respectively as "any activity undertaken by a person who believes himself to be healthy for the purpose of preventing disease or detecting disease in any asymptomatic state." A conceptual model called the Health Belief model designed to explain demand for these two types of visits has been proposed by Rosenstock (1966). The Health Belief model consists of two dimensions or consumer states. One is defined as "the psychological state of readiness to take specific action" and the other is defined as "the extent to which a particular course of action is believed, on the whole, to be beneficial in reducing the threat." Both cognitive and "emotional" psychological elements are viewed as affecting the individual's psychological readiness. These include perceived susceptibility to a disease or condition and its perceived seriousness if contracted. Rosenstock presents hypotheses indicating the action consumers will take given each state combination. The action hypotheses are, however, somewhat confusing and somewhat incomplete. Essentially, Rosenstock states that, when readiness is low, intense stimuli will be needed if action is to result. He also states that when readiness is high, even a slight stimulus will result in some action. But, because actions that may be effective in reducing the threat of illness may also be perceived as causing significant emotional, physical or financial discomfort, the consumer may simply avoid taking such action. The consumer might instead try to remove himself psychologically by engaging in some "activity" that he may be able to pretend is threat reducing, but which in actuality is not. Or, the consumer may experience an increase in fear, and as a result become incapable of taking reasoned action.

Feshbach (1953) study of fear arousing communications, but allowing for more diffuse responses to a state of high readiness. The model does not appear to be an unreasonable representation, but it does seem less precise than one would ideally like, because of the variety of responses it predicts. Ideally, one would want to know the circumstances under which the consumer will take the threat reducing action, or will indulge in psychological removal, or will take unreasoned action based on fear. Perhaps a conceptual scheme originally proposed to deal with situations such as those described by Janis and Feshbach (1953) can be of some assistance.

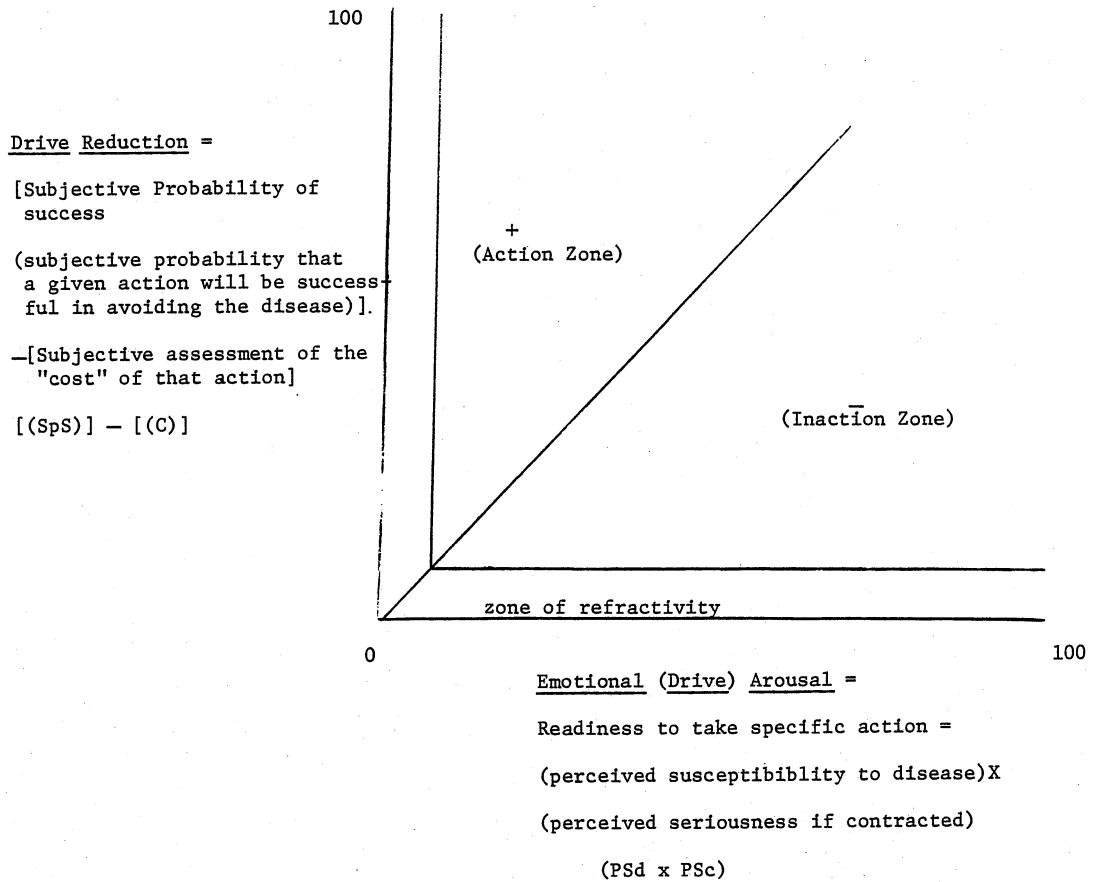
Bauer with Cox (1963) also working from findings such as Janis and Feshbach's, propose that communications can have two dimensions, an "emotional" or drive-arousing dimension, and a "rational," or drive-reducing dimension. They argue that the effect of a communication of a given strength of drive-arousing and/or drive-reducing content depends on the exact state of the recipient prior to receiving the communication. Exhibit 2 presents their argument diagrammatically.<sup>1</sup>

<sup>1</sup>There is no intention here to slight the most interesting model proposed by Zaltman and Vertinsky (1971). It is simply that the Zaltman-Vertinsky model was designed with an emphasis on less developed countries and is somewhat more elaborate than is necessary for the purpose of this paper.

Actually, the Rosenstock model is basically an extension and elaboration of the classic Janis and

Exhibit 2

Extended Formulation



The diagram is taken directly from Bauer with Cox (1963), but the labelling is an extension of the Rosenstock Health Belief Model, and the formulation to follow may include an idea or two of my own.

A line of unknown slope (here 45° is taken arbitrarily) divides the diagram into an action and an inaction zone. This line indicates that the more severe the threat (severity is the product of perceived susceptibility and perceived seriousness), the more "sure" the outcome of a threat reducing action must be in relation to its cost before the consumer will take that action. The exact probability of success required in order for a consumer to take any given action at any given level of readiness is a function also of the perceived cost of the action; at any given level of readiness, higher probabilities of success are required for high "cost" actions to be undertaken.

The conditions under which each of the alternative actions proposed by Rosenstock will be taken can now be hypothesized somewhat more precisely for consumers who are in a state of high readiness. These consumer actions are predicted on the basis of perceived subjective probability of success (SpS) and cost (C) of the action and they do not have to be predicated on reactions to a given communication. Specifically,

- a. Favorable (the "medically correct") action will result most readily when (SpS) is high and (C) is low.

- b. Inaction will result when all alternatives are perceived as low in both (SpS) and (C). The tendency toward inaction increases as (C) increases while (SpS) remains constant.
- c. An action that is perceived as threat-reducing but which in actuality is not will result when that action is perceived to have a lower (C) attached to it than to other alternative actions and when the additional (C) attached to "better" alternatives does not appear to be balanced by an increased (SpS).
- d. The consumer's response is simply increased fear when the seemingly "best" alternative has an unacceptably high (C) and/or a low (SpS). In this condition there is no alternative the consumer can pretend is threat reducing.

It should be a worthwhile exercise to look at some studies that report on consumer behavior with respect to preventive and therapeutic visits to see whether the hypotheses just stated seem to hold up. The first of such studies antedates the Rosenstock model but it is of interest both because it, itself, is a compilation and analysis of several other studies, and because of its subject matter:

It should be noted that in very extreme cases, e.g. where a consumer has been diagnosed as having an irreversibly fatal disease, it is likely that any and every possible action, regardless of its (C) or (SpS) might be undertaken.



poliomyelitis vaccination. Polio vaccination is purely preventive medicine. At the time the study to be reported here was compiled, Salk vaccine had been proven to be a highly effective polio preventive, and the vaccination procedure was a relatively easy one for the patient to undergo. Thus, the treatment had a high (SpS) and a low (C). Moreover, a minor epidemic of paralytic polio had just occurred, indicating that the study population might be expected to have been in a reasonably high state of readiness.

Rosenstock, Derryberry, and Carriger (1959) analyzed six studies on acceptance of polio vaccination which they considered methodologically sound. Their conclusions were that "a basic determinant of the decision... is the extent to which the individual believes that he is susceptible... or ... to which the parent believes ...his child is..." There seems to have been little question among consumers that the Salk vaccine was effective and could be administered with little discomfort to the patient. Non-vaccinatees, then, simply seemed to be consumers to whom polio was of little significance; data are cited which show that vaccination rates were higher in populations that (1) demonstrated stronger beliefs in the seriousness of the disease, or (2) were exposed to some form of "promotion" designed to raise the salience of the disease (thus increasing consumer's state of readiness).

There is no indication from any of the findings that raising the salience of polio had any effect other than to increase the consumer's propensity to undergo vaccination.

Studies of behavior in situations that have diagnostic as well as preventive overtones, however, appear to elicit consumer behavior that is somewhat more difficult to interpret. Fink, Shapiro and Roester (1972) describe an effort to increase participation in a breast cancer detection program. The effort was an experiment conducted over a four year period among a sample of female members aged 40-64 of the Health Insurance Plan of greater New York. Members in the experimental group were first contacted by a mailing which informed them of the initiation of a breast cancer detection study and asked them to make an appointment to participate in a screening examination. Non-respondents were sent a second letter reemphasizing the importance of the exam, and were further contacted by telephone if they did not respond to this second letter. Later contacts were made with each participant at annual re-examination time.

If one assumes that the contents of the communications were actually as described, the description indicates a communication that appears to be essentially drive reducing (more specifically, reducing the cost, or (C) term) since it appears simply to offer a convenient service. Thus, the communication should have been most effective with consumers who were already in a state of readiness. Study findings indicate that this was, in fact, the case. Participants in the first screening tended to be younger, to be more highly educated, more favorably disposed toward the concept of screening, and to express more concern about the possibility of having breast cancer than did non-participants.

Similar factors distinguish participants who completed all four exams from those who did not. Education, income and occupation were all positively correlated with completion. Both previous use of medical services and previously having had a polio vaccination were also positively related to completion. There was also a positive relationship between respondents' self-reported

state of health and their completion of the four screening exams.

Women who started the screening program but did not take all four examinations showed a significantly higher vote of agreement with the statement that "Physical examinations just make you worry; it's like looking for trouble" than did women who did take all four examinations. This finding would also be consistent with the model if, for these women, the examination itself could have been a drive-arousing rather than a drive-reducing experience. It is not difficult to imagine such a possibility; an uncommunicative or seemingly unsympathetic or uncaring physician can easily raise a patient's anxiety level, with the result that the patient will become a "drop out."

Given the high level of publicity relating smoking and disease, cigarette smokers ought to be in a perpetually high state of readiness with respect to medical problems. The risks involved in smoking might indicate more frequent exams. But, the extended model would predict that cigarette smokers should be less frequent users of preventive and diagnostic medical services than should non-smokers. This is because the diseases smokers contact as a result of their smoking are relatively incurable, and it's not unreasonable to expect that smokers know it.

Oakes et al (1974) studied participation in health examinations among a sample of Kaiser-Permanente Medical Care program members. Members prepay and are entitled to free physical examinations. A mailed questionnaire was used to elicit rates of participation in physical examinations. The study included procedures to adjust for non-respondents. Data were analyzed controlling for socio-economic status and for cigarette smoking. Among males, current smokers were less likely to have had any kind of health examination over the previous five years than were either ex-smokers or non-smokers. The difference was most pronounced among men of higher social class (who are likely to be best informed about the consequences). Among women, the differences between smokers and non-smokers were much less pronounced. It does not seem unreasonable to explain this finding by noting that women have other occasions, principally gynecological, for having regular physical examinations.

Predicting use of a center offering free, comprehensive medical care to a low income population for whom medical care had previously been inaccessible and unaffordable should also be an interesting exercise for the model. If such a center is well-promoted, it might even succeed in raising consumer's readiness status to too high a level. The result might be overuse of the center, even in circumstances where the consumer's (SpS) is relatively low. This is because certain dimensions of the consumer's (C) are also low: if the center is both free and easily accessible. Given that these (C) dimensions are low in every circumstance, overuse might be predicted specifically in those situations where the expected treatment is not uncomfortable.

A "before-and-during" study of consumer's attitudes toward the use of various medical care services in such a population was conducted among residents of a somewhat isolated Boston-area low income housing project (Bellin & Geiger, 1972). Two waves of interviews were conducted, both before and two years after the initiation of a comprehensive community health center in the project. Project residents were involved in the design and operation of the center; the center was heavily promoted and much talked about, and residents could use the center's facilities without charge.

Comparison of the results of the two studies showed that attitudes toward asymptomatic medical checkups become significantly more favorable and that the proportion of project residents who reported having had a general physical during the past 12 months rose from 17% to 59%.

This is certainly not an unreasonable finding, but it really says very little about the efficacy of the model. There is, however, another finding that is of some interest. Respondents in both waves were given a list of conditions or symptoms including several from each of three categories: serious (e.g. severe shortness of breath); mixed (e.g. feeling tired all the time); minor (e.g. sore throat, running nose). They were asked whether a doctor should be seen for each condition or symptom. In the baseline survey, a much larger percentage of respondents said "yes" for the serious conditions than for the minor ones. The percentage for the mixed conditions fell in between the serious and minor percentages. Results of the later survey showed a striking increase in the percentage of respondents who said "yes" for minor conditions (from 30% to 50% for sore throat, runny nose, for example). Again, these results are consistent with the model, and the thrust of the studies so far discussed suggest that the model can be a useful schema for working with preventive and diagnostic visits. Specifically, it suggests, schema for determining the content of promotional efforts that are based on consumers' existing attitude states. The model should be most effective in working with specific preventive and diagnostic problems.

Given any specific situation, straightforward research techniques can be used to establish existing consumer states with respect to (1) each of the readiness dimensions, (2) the perceived (SpS) of the desirable medical action, (3) perceived (SpS)'s of other medical activities that consumers might incorrectly perceive as desirable, and (4) the (C) for each possible medical action. Ideally, these data can be linked to other consumer characteristics, either demographic, or psychographic, or both, so that identifiable market segments can be defined.

These market segments can then be plotted on a graph such as the graph in Exhibit 2. Given the position of each market segment on the graph, it should be feasible (1) to identify the probable direction (up, down, right, left) that market segment would have to be moved in order for the medically favorable action to result and (2) to identify the way in which the content should be balanced for susceptibility, seriousness, (SpS), and (C) so that the medically favorable action would be most likely to result.

It should then be possible to construct specific promotional programs aimed at individual market segments, to test them and to measure the results. Both testing and measurement of results are often accomplishable since the marketer will usually be using either direct mail or other controllable media and/or personal selling efforts. In addition, consumers who take the medically favorable action can be specifically identified and "tied back" to specific promotional efforts and to specific market segments, and therefore to predisposition. Thus, the model's prediction can also be validated.

It is not likely, however, that the process just described will be quite as tidy in actuality as it appears on paper. For one thing, the actual slope of the line dividing the action and inaction zones in the model is unknown. For another, it may be difficult

to estimate exactly how strong a specific state of readiness may actually be. Even in the face of these complications, however, consumer behavior in preventive and diagnostic situations is likely to be somewhat more straightforward than behavior in therapeutic situations. It is to these situations that we shall turn.

#### The Decision to Seek Medical Care: Therapeutic Visits

Kasl and Cobb (1965) also provide a good description of therapeutic behavior, defining it as "any activity undertaken by a person who feels ill, for the purpose of defining the state of his health and of discovering suitable remedy." Any review of studies conducted on consumers' use of medical care services for therapeutic reasons quickly identifies that a visit to the physician may be the end, rather than the beginning of a behavior chain that was initiated when some symptom was experienced. And the behavior chain may end before the physician visit stage is reached.

A landmark study conducted among members of a Bronx, N.Y. prepaid medical plan (Friedson, 1961) discovered the existence of "lay referred systems" that consumers progressed through before they formally consulted a physician. Usually, the subject began the search for medical help with self-diagnosis and some resultant self-treatment. If the self-treatment did not appear successful, in some cases others would be consulted, first members of the household, and then outside lay people. In these cases, advice and diagnosis would be sought from each lay person, and the physician would be consulted only when the lay advisers' prescriptions proved ineffective.

These lay referral systems, however, appeared to differ based on the social class of the patient, because patients of upper versus lower social classes evidenced different views of their bodies and bodily processes, and of illness. Essentially, lower class people tended to be apprehensive about illness and to be relatively ignorant with respect to their bodies and bodily functions. They also evidenced considerable ignorance about illness, and about types of medical treatment. They tended to view the physician as something of a mystic, whose skills were arcane, incomprehensible, and above questioning. The upper class patient, on the other hand had a much more realistic and detached view of his or her body and its functions, and evidenced much more sophistication with respect to illness. This patient tended to view the physician much more as a colleague, and to take a somewhat detached but realistic and concerned interest in the treatment the physician was providing.

The lower social class referral system tended to be localized to kin and neighbors, selected on the basis of family relationship and physical proximity. The upper class lay referral system did not depend on kin or neighbors; advisers were selected on the basis of perceived knowledge, rather than on the basis of family or neighborhood. A summary view of these social class differences might describe lower social class patients as unwilling to try the system except as a last resort, but blindly trusting the system once they entered it. It might describe upper social class patients as much more willing to try the system, but also much less trusting of the system once they enter it.

Suchman (1965) conducted an extensive study, also among New York residents, that supported many of Friedson's

(1961) findings. He hypothesized that an individual's health behavior resulted from that individual's medical orientation, and that the individual's medical orientation was a consequence of his or her socio-demographic characteristics and thus the concomitant community orientations. His findings identified social class linked community systems that paralleled the "lay referral" systems of Friedson (1961). Lower social class systems tended to be parochial, while higher class systems tended to be more cosmopolitan. Lower social classes also evidenced "lower knowledge about disease and higher skepticism toward medical care..." which again supports Friedson's (1961) findings. The thrust of these findings would suggest that social class membership plays an important role in determining the therapeutic aspect of medical care utilization; however, Suchman's data are not strong where the link between medical orientation and behavior must be made. At least one, more recent study fails to replicate all of the Friedson (1961) and Suchman (1965) findings.

Reeder and Berkanovic (1973) conducted a study in the Los Angeles area in 1970 using many of Suchman's variables. They did not find a significant relationship between social class and medical orientations. These researchers cited all of the usual reasons for the failure to replicate, the most compelling of which seem to be time and place. The late 1960's, and California may well have bred a lower social class population that is more aware of good medical care and the medical care system. As the Bellin and Geiger (1972) study cited earlier suggests, perhaps people can learn to use the medical care system for therapeutic reasons and there may now be other determinants of its use such as heightened feelings of entitlement, that are more important than social class.

In any event, the conclusion of a review of several studies conducted among different populations (Zola, 1972) and concerned with different medical problems should not be surprising. This review concluded that a substantial number of medical complaints are not brought to physicians for treatment (at least they are not purposefully and specifically brought to physicians). However, as Zola (1972) forcefully points out, this does not mean that these complaints go completely untreated. Self-medication (or medication prescribed by a friend or pharmacist), a visit to a paramedical (podiatrist, osteopath etc.) and "informal" medical consultation (while the physician is treating another family member, for example) are among the many ways in which some of these medical conditions actually do get treated.

Whether such behavior is functional or not depends both on the point of view of the analyst and on the specific medical conditions involved. For example, if the analyst is a physician in a fee-for-service practice, great concern might be expressed that any complaint may be going untreated. A physician in an organization offering a prepayment plan, however, could consider the situation highly functional, depending on the specific medical condition that may be going without physician treatment. Both, however, might be interested in shaping behavior, although toward somewhat different ends. Thus, it should be useful to explore (1) what complaints tend to go without treatment, and, (2) whether there are systematic differences among consumers with respect to what conditions do or do not get treated.

Ludwig and Gibson (1969) studied the extent to which medical care was sought among a population of appli-

cants for social security disability benefits who claimed to be in "poor" health or worse. He divided this population into two groups, those who had "seen a doctor or visited a clinic or hospital" within two months (users), and those who had not done so for six months or more (non-users). Data were collected for both groups on their recognition of symptoms, situation factors affecting the ease or difficulty with which medical attention could be obtained, and faith in the medical system. The results of the study are quite interesting. The type and number of symptoms respondents reported had no significant effect on their use of medical care; respondents with any specific symptom were no more or no less likely to have sought medical care than were respondents with any other specific symptom. Respondents reporting 5 or more symptoms were no more likely to report having sought medical care than were respondents with 4 or fewer symptoms. Both faith in the medical care and the ease of obtaining medical attention, however, affected the use of medical care. Among respondents who had a positive medical-scientific orientation (as measured by a three-item index), just 7% had not sought medical care in the past 6 months, while 32% of respondents with negative orientations had not done so.

There was a strong positive relationship between income and use of medical care: 37% of very low income respondents versus just 10% of somewhat higher income respondents had not sought care. Recent welfare contact inhibited medical care seeking: respondents who had recent contact with welfare agencies were twice as likely not to have sought medical care. Among this study population, therefore, it was not a question of what symptoms went untreated as much as a question of what beliefs and circumstances facilitated or inhibited the seeking of care. It should also be noted that this study deals with an essentially a lower social class population, yet both positive and negative medical-scientific orientations were found. In addition, some other variables appeared to be at least as important as respondents' medical-scientific orientations in predicting their use of the medical care system.

Hackett, Cassem and Raker (1973) provide strong further reinforcement for the finding that it is not primarily the presence or absence of symptoms that causes consumers to use medical services. They surveyed both in-and-out patients visiting the tumor unit at the Massachusetts General Hospital during the period 1968-70. Their survey collected data on delay (the time between each patient's first awareness of a sign or symptom and the first physician visit). Also collected were data on each patient's reason for first visiting the tumor unit, a self-description of their conditions (after diagnosis), a self-rating of delay time, the extent to which they worried about their health, family cancer history, and social class.

The researchers found considerable evidence of delay among study respondents. Just 20% had consulted a physician within a week of becoming aware of the first symptom or sign, and 66% had waited at least a month. The least amount of delay was found among patients whose reason for visiting the tumor unit was that their tumors were discovered during a routine physical examination. Those whose reason was worry about their symptoms showed the next to least amount of delay those who came in because their symptoms were incapacitating had delayed still more, while those who visit the tumor unit because of pain had delayed the longest. Delayers also tended to deny their condition, often calling it a "tumor," while those who were more prompt in seeking treatment were more likely than delayers to refer to their condition

as "cancer." Those who worried about their health delayed somewhat, but not significantly more than those who did not. Respondents who reported a family history of cancer delayed somewhat longer than those who did not report such a history, and those who claimed they generally procrastinated in seeing physicians were more likely to have delayed in this particular instance. Patients higher in social class tended to delay less than lower social class patients; however, this finding may simply be related to a greater propensity of people in higher social classes to have regular physical exams.

It is difficult to disagree with these researchers' conclusion that the patients who delayed did so knowingly and proposefully. It is also difficult not to look at the findings from this study in light of the model presented in Exhibit 2. Obviously, the presence of signs or symptoms indicates a consumer in a high state of readiness, which suggests that the way to encourage a physician visit is through drive reduction. The group who visited the tumor unit with the least amount of delay were those whose signs or symptoms were discovered during a routine physical examination. These are also the patients who were the most likely to have been exposed to a drive reducing communication, since the physician could have explained the relationship between early treatment and success. Unfortunately, this study did not measure respondents' perceptions of the likelihood of successful treatment.

#### Emerging Generalizations

This paper has attempted to present some findings from a selected group of studies covering a range of consumers and medical situations. In general, it would appear that medical care consumer behavior is shaped more by psychographic than by demographic factors. Where demographics appear to predict behavior, it may well be that these demographic factors correlate closely with associated psychographic factors which are more likely to be causal. Such a generalization, of course, is consonant with expectations derived from the study of consumer behavior in the commercial sector.

However, the thrust of these findings may also be at cross currents with commercial sector findings that point to an increasing regard for the role of situation specific factors in shaping behavior. With respect to health care behavior, we might be more inclined to suspect a basic orientation toward medicine and medical care that is significant in predicting behavior in any given situation. This basic orientation might predict most effectively where there are no institutional or other barriers to action. That is to say, a given consumer's behavior toward a polio shot versus a cancer screening examination, versus a symptom for example might be more uniform than that same consumer's behavior in purchasing a detergent versus a set of tires. However, such a speculation must be made with caution if one's concern is with producing behavior change.

The studies reviewed also suggest that behavior can be changed in specific health care situations. It would be worthwhile to know the extent to which changing the consumer's behavior in any given situation has a carry-over effect in terms of changed behavior toward other health care situations. Such knowledge could be extremely useful, for example, in finding the most efficient entry points to produce the most pervasive health behavioral change.

#### Bibliography

- Raymond A. Bauer with Donald F. Cox, "Rational Vs. Emotional Communications: A New Approach," in L. Arons and M. May, eds. Television and Human Behavior, Appleton-Century-Crofts, New York, 1963.
- Eunice C. Beale and Steven A. Schroeder, "Marketing for an Urban Health Center," Health Services Reports, Volume 8, 1973 pp. 84-88.
- Seymour S. Bellin and H. Jack Geiger, "The Impact of a Neighborhood Health Center on Patients' Behavior and Attitudes Relating to Health Care," Medical Care, Volume 10, No. 3, May-June 1972 pp. 224-239.
- Bionetics Research Laboratories, Inc., Marketing of Health Maintenance Organization Services, U.S. Dept. of H.E.W., Washington, D.C., 1973.
- Raymond Fink, Sam Shapiro, and Ruth Roester, "Impact of Efforts to Increase Participation in Repetitive Screening for Early Breast Cancer Detection," American Journal of Public Health, March 1972, pp. 328-336.
- Eliot Friedson, Patients' View of Medical Practice, Russell Sage Foundation, New York, 1961.
- Thomas P. Hackett, N.H. Cassem, and John W. Raker, "Patient Delay in Cancer," The New England Journal of Medicine, Volume 289, No. 1, pp. 14-20.
- Irving L. Janis and Sidney Feshbach, "Effects of Fear Arousing Communications," J. Abnormal & Social Psychology, 48, 1953, pp. 78-92.
- Stanislav V. Kasl and Sidney Cobb, Health Behavior, Illness Behavior, and Sick Role Behavior, ISR, University of Michigan, Ann Arbor, 1965.
- Philip Kotler, Marketing for Nonprofit Organizations, Prentice Hall, Inc., N.J., 1975.
- Edward G. Ludwig and Geoffrey Gibson, "Self Perception of Sickness and The Seeking of Medical Care," Journal of Health and Social Behavior, II, July 1969, pp. 125-133.
- Thomas W. Oakes et al, "Health Service Utilization by Smokers and Nonsmokers," Medical Care, Volume 12, No. 11, November, 1974, pp. 958-966.
- Leo G. Reeder and Emil Berkanovic, "Sociological Concomitants of Health Orientations: A Partial Replication of Suchman," Journal of Health & Social Behavior 14, June 1973, pp. 134-143.
- Irwin M. Rosenstock, Mayhew Derryberry, and Barbara K. Carriger, Public Health Reports, Volume 74, No. 2, February 1959, pp. 98-102.
- Irving M. Rosenstock, "Why People Use Health Services," Milbank Memorial Fund Quarterly, Volume 44, 1966, pp. 94-127.
- Edward A. Suchman, "Social Patterns of Illness and Medical Care," Journal of Health and Human Behavior, 6, Spring 1965, pp. 2-16.
- Gerald Zaltman and Ilan Vertinsky, "Health Service Marketing: A Suggested Model," Journal of Marketing, Volume 35, July 1971, pp. 19-27.
- I.K. Zola, "Studying the Decision to See a Doctor," Adv. Psychosom. Med., Volume 8, 1972 pp. 216-236.

## THE ROLE OF MASS COMMUNICATIONS IN PROMOTING PUBLIC HEALTH

Mary Jane Schlinger,  
University of Illinois at Chicago Circle

### Abstract

The beliefs, attitudes and motives that underlie poor health practices often are deeply-rooted and emotion-laden. For this and other reasons, it typically is difficult to persuade people to engage in activities which relate to health maintenance. Recently published literature indicates that there is some disillusionment with mass media campaigns as a tool for changing health behavior. Although such campaigns can make an impact, they often do not. The present paper discusses the role of mass communications in health marketing programs and suggests some ways in which the planning, implementation and evaluation of media efforts can be improved.

### Introduction

In 1969, Kotler and Levy published a seminal article, "Broadening the Concept of Marketing," which stimulated widespread interest in applying the principles and concepts of marketing outside of the traditional for-profit business arena. Concomitantly, marketing researchers began focusing increased attention on the consumer decision-making process as it operates in such diverse areas as politics, religion, education and health.

The entire issue of the July, 1971, Journal of Marketing was devoted to what might be called public service or social marketing, i.e., the marketing of ideas, services, products and organizations that are intended to help accomplish societal goals. The 1971, 1972 and 1973 Association for Consumer Research national conferences each incorporated sessions on broadening the concept of consumer behavior. The following year, a series of papers was delivered on marketing political candidates. And in 1975, the Association for Consumer Research program included a session titled "Consumer Behavior and Public Health."

The purpose of this paper is to examine the "Consumer Behavior and Public Health" presentations and other selected literature for implications about the potential role of the mass media in promoting public health. Before doing that, however, it would be useful to review some of the problems that face would-be marketers of public health.

#### Problems Faced by Public Health Marketers

The marketer of programs, ideas, services and products designed to improve public health faces a formidable task.

For one thing, public health organizations typically aim at eliciting some kind of short- or long-term behavioral response from large numbers of people, often 50% to 100% of a population or population segment (Rosenstock, 1960). In contrast, a ready-to-eat cereal manufacturer probably would be happy, or even ecstatic, if a new brand captured a four per cent share of the cereal market, or if an existing brand improved its sales by a one per cent share.

Not only do health agencies frequently need to reach large audiences, they often also must deal with the "laggards" and "no nothings," i.e., that fifteen per cent or

so of the population which is poorly educated, ill-informed, suspicious of innovations, traditional in outlook, and backward rather than forward in orientation (Hyman and Sheatsley, 1947; Rogers and Shoemaker, 1971).

For another thing, public health workers face a paradoxical situation in which the beliefs, attitudes and motives which underlie poor health practices may be deeply-rooted, intransigent and emotion-laden, whereas the beliefs, attitudes and motives which support good health may be absent or lacking in saliency (Rosenstock, 1960).

Smoking, high cholesterol diets and inadequate exercise are examples of unhealthy habits which can be the firmly established results of years of learning and reinforcement.

Self breast examination is an example of a health-related measure which is suffused with emotion and fear. A Gallup poll conducted in late 1973 found that few women even claimed to examine their breasts on any regular basis. In part, this may be due to ignorance of the importance of self-examination, lack of knowledge or self-confidence about performing the examination, and uncertainty about possible outcomes if a lump is discovered. But perhaps even more important is the fear of cancer, a disease so dreaded that the word "cancer" once was taboo in polite conversation and in many mass media vehicles. Equally important is the emotion associated with the breast as a symbol of sexuality, femininity and motherhood (Reneker and Cutler, 1952; Goldson, Gerhardt and Handy, 1963). Because of this symbolism, women are perhaps inordinately terrified of mastectomy, and self-examination of the breasts may carry anxiety-producing overtones of masturbation.

Turning to the other side of the coin, that is, the beliefs, attitudes, motives and actions which support health maintenance and disease prevention, it appears that in the absence of overt symptoms many people feel safe or even invulnerable from disease; health concerns are forgotten, suppressed or repressed, interest in the subject can be relatively low, and there may be reluctance to interrupt established daily routines for any but the simplest additional health maintenance measures. Thus the health marketer has three options: (1) to increase the salience of the need for good health, as well as interest in maintaining good health; (2) to appeal to needs that are unrelated to health per se but which will lead to health-related behavior; or (3) to "package" the product or idea with other products or ideas that are desired by the target market. For low income consumers, for example, information on nutrition might be combined with information on how to buy and prepare food economically, assuming that saving money in the grocery store has higher consumer priority than nutrition.

Still another problem facing the public health marketer is the actual and psychological distance that often stands between the consumer and the product, idea or service (Wiebe, 1951-52; Rosenstock, 1960). Marketers of consumer goods try to make it easy, simple, convenient and even pleasurable to obtain their offerings. The user of public health services may have to travel long distances to overcrowded facilities, interact with unfamiliar and perhaps hostile persons, and submit to procedures that are inconvenient, physically or psychologically painful, time-con-

suming, difficult to understand, and expensive in terms of time, money and effort. Furthermore, the reward for many health-maintenance and disease-prevention actions is intangible, delayed or uncertain.

One final reason why it can be difficult to generate acceptance of public health programs and ideas lies in the nature of the profession and of some health organizations. Custom, ethical codes, budgetary considerations, bureaucratic red tape and sometimes laws prevent many health organizations from most effectively manipulating the four marketing variables: price, distribution, product and promotion. Also, public health professionals who know a great deal about epidemiology, microbiology and other such matters often know little or nothing about modern marketing.

The preceding discussion is not intended to suggest that the principles and practices of marketing as developed by businesses do not apply to the marketing of public health ideas, services and programs. In fact, certain commercial products may share one or more of the marketing problems delineated in the above paragraphs. The point is that the "marketing mix" -- the particular combination of a product, a price, a distribution system, and a promotional program -- must be appropriate to the nature of the public health marketing task.

The remainder of this paper will focus on one aspect of a marketing program: mass media promotion and publicity.

#### Do Mass Communication Campaigns Work?

Recently published literature suggests that there is some disillusionment with mass media advertising as a tool for changing health behavior. Haskins (1968) surveyed the literature on the effects of mass communications on drinking while driving and concluded that because most programs either had not been evaluated or had been evaluated inadequately, no one knows whether campaigns on drinking and driving have had any effect. A controlled experiment designed to evaluate an advertising campaign aimed at increasing seat belt usage showed that the promotional effort had no effect on belt use at all (Robertson et. al, 1974). A television campaign urging Illinois employers to hire the disadvantaged actually reduced employers' stated intentions to do so (Haefner, 1975).

In contrast, Maccoby and Farquhar (1975) and Maccoby (1975) have reported on an intensive two-year mass media campaign which was successful in reducing risk factors associated with heart disease among the target population. However, the behavioral changes brought about by the mass media campaign alone were very small compared to the changes generated by an experimental condition which combined mass media advertising with face-to-face group instruction for high-risk subjects.

Thus it appears that although mass media campaigns can make an impact on health behavior, they often do not. Disappointment with mass media efforts in the past frequently can be attributed to unrealistic expectations on the part of campaign planners as to what a given effort can accomplish. The real question, then, is not "Do mass communication campaigns work?", but, more appropriately, "How do mass media efforts work and what can they achieve?"

#### What Can Mass Communications Campaigns Do?

Existing evidence indicates that mass media efforts to

improve public health can and have accomplished the following tasks: (1) increase awareness of a health problem; (2) raise the level of information about health topics; (3) make a health topic or problem more salient, thereby sensitizing the audience to other efforts, such as personal selling, group education or direct-mail brochures; (4) stimulate interpersonal influence via conversations with family, friends, doctors and other experts; (5) generate forms of self-initiated information seeking; and (6) reinforce existing attitudes and behaviors.

Of course, given enough time and money, mass media campaigns occasionally may operate directly to change some health related behavior, such as product advertising can successfully pre-sell certain low-priced consumer grocery and drug items, given advertising budgets that are large enough to achieve adequate levels of exposure. But the point is that there are a number of other possible outcomes of public health advertising and publicity.

#### Planning, Implementing, and Evaluating Public Health Campaigns

Recent efforts at utilizing the mass media to promote health implicitly or explicitly point to some strategies for improving the effectiveness of promotional campaigns, and these will be reviewed in the following paragraphs.

#### Planning Campaigns

A number of observers have emphasized the importance of setting explicit and specific goals for mass media campaigns (Mendelsohn, 1973; Ray and Ward, 1975). As with any promotional goals, those in the health area must state what target market or markets are to be reached, what message or messages are to be delivered, which media are to be used, and what measurable effects are to be achieved over what time period. The statement of prior goals is important for several reasons. First, the goals provide guidance for copywriting and media selection. Ward (1975) has noted that unless there are explicitly stated goals advertising copywriters may treat public service campaigns like "creative trips," where they are free to let their imaginations run wild without the pressures of accountability that are exerted by commercial clients. Second, the statement of promotional goals guides evaluation, permits evaluative plans to be specified in advance, and allows for the collection of baseline data where that is desirable. And third, the process of stating promotional objectives should aid the campaign planner in estimating budgetary needs and in keeping the objectives realistic, given resource constraints.

Another important element in planning successful health campaigns is pre-research on the target market. Extrapolating from Rosenstock's (1966) widely-known health belief model and research by Bauer and Cox (1963), Wortzel (1975) suggests that the health marketer should study target markets in terms of their perceptions with regard to: (1) their susceptibility to a health disease or problem; (2) the seriousness of the disease or problem if it is encountered; (3) the estimated probability that a given course of action will reduce the threat; and (4) the estimated cost of the action. Wortzel believes that data on the above factors can be combined with other demographic and psychographic information in order to define market segments and select appeals for each that are most likely to motivate behavior.

After systematically reviewing two decades of literature on the utilization of health and welfare services,

McKinlay (1972) concluded that there is a need for more small-scale exploratory studies designed to generate hypotheses about consumers, and that insufficient attention has been paid to the influence of kin and friendship networks on health and illness behavior.

The point is that mass media campaigns are more likely to be effective if they are based on an understanding of relevant demographic, psychological, and sociological characteristics and experiences of the target consumers. And since consumer behavior literature within the public health field is not ordinarily geared to practical marketing planning, it is quite probable that literature reviews will need to be supplemented with primary research.

#### Implementing Campaigns

One of the problems faced by any public service marketer is obtaining high-quality, professional creative work. One solution is to ask advertising agencies to donate creative services; many agencies do donate time to worthy causes either through the Advertising Council, local advertising organizations, or independently. The client, of course, must pay the costs of production. However, advertising agencies are not likely to provide the same quantity or quality of time, talent and collateral services to non-paying public service clients that they provide to paying clients, and it is only realistic to recognize that fact. Also, there is the previously mentioned danger of unsupervised copywriters indulging in "creative trips" on public service projects.

Both Mendelsohn (1973) and Maccoby (1975) argue that the practitioners who create advertising must understand and be guided by empirically defined principles of communication and persuasion and not allowed to proceed on the basis of subjective insight, guesswork and intuition. In fact, Maccoby believes that an important strength of the previously mentioned Stanford campaign to reduce heart risk factors is that the media campaign was designed, created and pretested by the University's Department of Communication media staff which was trained in communications research.

Many observers agree that copytesting is a key to successful public service advertising. Maccoby (1975) has noted that the advertisements and other materials utilized in the Stanford campaign to "unsell" heart disease were extensively pretested. And Swineheart (1975a, 1975b) reported that the Children's Television Workshop conducted or commissioned some sixty formative research studies in the course of creating the Public Broadcasting System's health program, "Feeling Good."

The importance of copytesting public service advertising is particularly apparent where that advertising addresses emotional or controversial issues, since there may be a substantial danger of viewer distortion. For example, copytests of a number of storyboards designed to promote ethnic plurality and pride in national and racial origin indicated that one of the storyboards delivered exactly the opposite message, i.e., that everyone should be alike so there would be no prejudice, and another storyboard, which intended to portray and then refute prejudiced beliefs, failed to achieve an effective refutation (Leo Burnett U.S.A., 1974).

Ray and Ward (1975) provide an overview of the issues involved in copy pretesting, and suggest a multifaceted set of copytesting procedures.

#### Evaluating Campaigns

Where properly controlled experiments are feasible, it is important to evaluate promotional efforts in terms

of behavioral change. However, a number of researchers offer persuasive arguments for examining the cognitive and affective as well as the behavioral influences of promotion (Clarke and Kline, 1974; Kline, Miller and Morrison, 1975; Ray and Ward, 1975). The values of a multifaceted approach to the evaluation of mass media campaigns are several: (1) if only behavioral change is measured, many effects of the campaign will be overlooked; (2) it will be difficult to understand why or how a campaign succeeded or failed to modify behavior if other responses are not measured; and (3) it is valuable to see how cognitive, affective and behavioral responses relate to one another and to demographic and nondemographic characteristics of the respondents. Swineheart (1975) notes the importance of assessing exposure levels to mass media campaigns, since those serve as a constraint on further effects. Robertson (1975) emphasizes the significance of observing certain kinds of behavioral outcomes, such as seat belt use, where verbal questions may generate overclaiming. And Ray and Ward (1975), Clarke and Kline (1974), and Douglas, Westley and Chafee (1970) point out that evaluative studies should not overlook the influence of mass media on subsequent interpersonal communication.

#### Conclusion

This paper has reviewed recent literature on marketing public health programs and ideas to assess some of the implications for planning, implementing and evaluating mass media campaigns aimed at improving health. It appears that some health professionals may hold an exaggerated view of the mass media's capability for modifying health-related behavior. The actual effects of advertising on deeply-rooted behavior is apt to be relatively small at best, and to require careful planning and substantial resources. The implications of this are several: (1) mass media campaigns in public health may be best suited to the achievement of intermediate goals, such as awareness, increased salience of a topic, etc.; (2) where behavioral change is the goal, it may be necessary to combine mass media with interpersonal and group re-education programs; (3) there is a need to experiment with nontraditional formats for teaching new behavior via the media, perhaps by simulating interpersonal and group learning sessions in some way; and (4) in many cases, such as seat belt usage, where behavior is particularly intractable, persuasive strategies cannot achieve public health goals and other solutions, such as legal sanctions, money incentives, etc., must be sought.

#### References

- Raymond Bauer and Donald F. Cox, "Rational Versus Emotional Communication: A New Approach," in Leon Arons and Mark May, eds. Television and Human Behavior. New York: Appleton-Century-Crofts, 1963, 140-54.
- Peter Clarke and F. Gerald Kline, "Media Effects Reconsidered: Some New Strategies for Communication Research," Communication Research, 1 (April, 1974), 224-40.
- "Communication Study of Six Storyboards." Chicago: Leo Burnett U.S.A., 1974. (author same as publisher)
- Dorothy F. Douglas, Bruce H. Westley, and Steven H. Chafee, "An Information Campaign That Changed Community Attitudes," Journalism Quarterly, 47 (Autumn, 1970), 479-87, 492.
- Rose K. Goldsen, Paul Gerhardt, and Vincent Handy, "Patient Delay in Seeking Cancer Treatment: Behavioral Aspects," Journal of Chronic Diseases, 16 (1963), 427-36.

James E. Haefner, "Television Advertising: A Study on its Applicability for Promoting the Disadvantaged to Employers," in Mary Jane Schlinger, ed., Advances in Consumer Research, Vol. 2. Chicago: Association for Consumer Research, 1975.

Jack B. Haskins, "Evaluation of Effects of Educational and Mass Communication Techniques" Transactions. National Safety Congress, 24, (1968), 77-109.

Herbert Hyman and Paul B. Sheatsley, "Some Reasons Why Information Campaigns Fail," Public Opinion Quarterly, 11, (1947), 412-23.

F. Gerald Kline, Peter V. Miller, and Andrew J. Morrison, "Communication Issues in Different Public Health Areas," in B.B. Anderson, ed., Advances in Consumer Research, Vol. 3. Cincinnati, Ohio: Association for Consumer Research, in press.

Philip Kotler and Sidney Levy, "Broadening the Concept of Marketing," Journal of Marketing, 33 (January, 1969), 10-15.

Nathan Maccoby, "Communication in Disease Prevention," presentation at the national conference of the Association for Consumer Research, Cincinnati, Ohio, 1975.

Nathan Maccoby and John W. Farquhar, "Communicating for Health: Unselling Heart Disease," Journal of Communication, 25 (Summer, 1975), 114-26.

John B. McKinlay, "Some Approaches and Problems in the Study of the Use of Services - An Overview," Journal of Health and Social Behavior, 13 (June, 1972), 115-52.

Harold Mendelsohn, "Some Reasons Why Information Campaigns Can Succeed," Public Opinion Quarterly, 37 (Spring, 1973), 50-61.

Michael L. Ray and Scott Ward, "Experimentation for Pre-testing Public Health Programs: The Case of the Anti-Drug Abuse Campaigns," in B.B. Anderson, ed., Advances in Consumer Research, Vol. 3. Cincinnati, Ohio: Association for Consumer Research, in press.

R. Renneker and M. Cutler, "Psychological Problems of Adjustment to Cancer of the Breast," Journal of the American Medical Association, 148 (1952), 833-8.

Leon S. Robertson, Albert B. Kelley, Brian O'Neill, Charles W. Wixom, Richard S. Eisworth, and William Hadson, Jr., "A Controlled Study of the Effect of Television Messages on Safety Belt Use," American Journal of Public Health, 64 (November, 1974), 1071-1080.

Everett M. Rogers and F. Floyd Shoemaker, Communication of Innovations (New York: The Free Press, 1971).

Irwin M. Rosenstock, "What Research in Motivation Suggests for Public Health," American Journal of Public Health, 50 (March, 1960), 295-302.

Irwin M. Rosenstock, "Why People Use Health Services," Milbank Memorial Fund Quarterly, 44 (1966), 94-127.

James W. Swinehart, "Creative Use of Mass Media to Affect Health Behavior," paper presented at the National Cancer Institute Conference on Cancer Control and the Behavioral Sciences, January, 1975a.

James W. Swinehart, "Developmental and Evaluative Research for a Television Series on Health," presentation at the national conference of the Association for Consumer Research, Cincinnati, Ohio, 1975b.

Scott Ward, letter to A. Sawyer dated October 22, 1975.

G.D. Wiebe, "Merchandising Commodities and Citizenship on Television," Public Opinion Quarterly, 15 (Winter, 1951), 679-91.

Lawrence H. Wortzel, "The Behavior of the Health Care Consumer: A Selective Review," in B.B. Anderson, ed., Advances in Consumer Research, Vol. 3. Cincinnati, Ohio: Association for Consumer Research, in press.



## PRE-PURCHASE INFORMATION ACQUISITION:

### DESCRIPTION OF A PROCESS METHODOLOGY, RESEARCH PARADIGM, AND PILOT INVESTIGATION<sup>1</sup>

Jacob Jacoby,  
Robert W. Chestnut (student), Karl C. Weigl (student), and William Fisher (student)  
Purdue University

#### Abstract

Using a methodology which captures consumer information acquisition as a dynamic behavioral process, four relatively clear-cut search strategies were identified: brand name reliance; predominantly within-brand search; predominantly within-dimension search; and complex search.<sup>2</sup> Path analysis suggests search strategy to be partially determined by brand loyalty and consumption frequency. These findings are considered in terms of a general "coding" explanation and placed within a four-stage paradigm developed for guiding, integrating and interpreting consumer information processing research.

#### Introduction

The three-fold purpose of this paper is to describe: (1) a methodology developed for studying consumer information acquisition as a dynamic process; (2) a paradigm for guiding research and integrating findings from consumer information processing investigations, and (3) a pilot study which employed this methodology and paradigm. Our objective in so structuring this paper is to have it serve as a focal statement around which a series of completed, on-going, and planned studies can be positioned and interpreted.

#### A Process Methodology for Studying Consumer Information Acquisition

##### Background

The present methodology evolved from a programmatic series of investigations which, at the outset, examined the implications of information load in the consumer context. By 1974, we had conducted more than 20 investigations directed at different aspects of consumer information processing. As detailed elsewhere (Jacobcy, 1975), despite our extensions and improvements, we became increasingly aware of limitations in our approach. Three things in particular disturbed us. First, most

<sup>1</sup>Preparation of this report was facilitated, in part, by a grant from the National Science Foundation (GI-43687) to the senior author.

<sup>2</sup>A variety of descriptive labels have been employed in recent attempts to categorize information acquisition strategies. Most often used are the terms "within-brand" and "within-dimension." To reduce confusion, the following can be equated. Interdimensional search (Payne, 1975), choice by processing brands (Bettman and Jacobcy, 1975), type "2" search (this paper), and within-brand search all denote the successive acquisition of some set of information dimensions on a single brand or alternative. In contrast, intradimensional search (Payne, 1975), choice by processing attributes (Bettman and Jacobcy, 1975), type "3" search (this paper), and within-dimension search all denote the successive comparison of some set of brands or alternatives on a single dimension of information.

of our early studies employed hypothetical rather than actual brands and product information. Second, operating within the tradition of manipulating levels of variables in tight experimental designs, subjects were provided with fixed amounts of product information (e.g., 12 items of information for each of 8 hypothetical brands) and required to arrive at purchase decisions based only upon the information provided -- no more, no less. In contrast, subjects in actual shopping situations are free to select as much or as little information as they wish prior to making purchase decisions. Finally, the designs employed were basically static in nature, focussing on the outcome of consumer decision processes rather than upon the dynamic nature and content of the process itself.

By late 1972, while our main research thrust continued to reflect these problems, we began developing more satisfactory process oriented approaches. Two such investigations were conducted during the 1972-73 academic year (Berning and Jacobcy, 1974; Jacobcy, Szybillo, and Busato-Schach, 1974). Further methodological developments ensued and culminated in a proposal submitted to the National Science Foundation in mid-1973.

This proposal was funded effective June 1974 -- with a novel twist. In an attempt to maximize external validity and practical utility, a special Research Advisory Board was established in conjunction with the grant and was composed so that various interests were represented thereon. James Turner (of Swankin & Turner), a prominent public interest attorney in Washington, represented the public sector. The business sector was represented by Norman Pugh of Sears, Roebuck & Co. Government interests were represented by Richard Herzog of the Federal Trade Commission, Evelyn Gordon of the Food and Drug Administration, and Alden Manchester of the U.S. Department of Agriculture. William Wilkie, a marketing professor who had worked at the FTC and was familiar with public policy issues relating to consumer information processing (cf. Wilkie, 1974; Wilkie and Gardner, 1974), and Howard Fromkin, a social psychologist known for his expertise in behavioral science research methodology (cf. Fromkin and Streufert, 1976), represented the academic perspective. George Brosseau, of the National Science Foundation, participated as an ex officio member.

The Research Advisory Board convened on July 11, 1974 and devoted most of the day to discussing the methodology described in the proposal as well as several alternative procedures. The most desirable features of each were fused into a single approach which, in general, represented an integration of the Berning and Jacobcy (1974) and Jacobcy, Szybillo, and Busato-Schach (1974) methodologies.

##### Methodology

Within the context of our program, several different process methodologies have been developed and used to examine a variety of consumer information processing issues (cf. Jacobcy, 1975). In all instances, we have focussed on pre-purchase information acquisition which, it should be noted, can and often does differ in nature and

content from pre-consumption and predisposition information acquisition (cf. Jacoby, Berning, and Dietvorst, 1975). For example, price information may be used prior to making a purchase, but be ignored prior to usage of the product. Suggested oven temperature may be ignored during pre-purchase deliberations, but be of vital importance when the consumer begins to cook. Whether an empty aerosol container can be safely placed in an incinerator might be irrelevant during pre-purchase and pre-usage information acquisition but become an important item of information prior to disposal.

On a conceptual level, the most basic form of the methodology involves presenting the subject with an array of product information and permitting him to acquire as much or as little of this information as he wishes prior to arriving at a purchase decision. The actual form that this array may take can and does vary. In the investigation to be described below, as well as in several other investigations now in progress, the format adopted consists of presenting information via a specially designed information display board (IDB) which presents information arranged in the form of a  $p \times q$  matrix. The various brands (i.e., purchase alternatives) appear listed across the top of the matrix while the types of information available (i.e., information dimensions, e.g., price, net weight, ingredients, etc.) are listed down the left- and right-hand sides of the matrix. Thus, each column of information in the matrix represents a single brand while each row represents a dimension of information (e.g., price) across all the brands presented. Each cell in the matrix contains the actual "value" of the information (cf. Jacoby, 1974, p. 117-18 and Jacoby, 1975, p. 206, for a more detailed discussion of this nomenclature) for the particular brand  $\times$  information dimension crossing which forms that cell. For example, crossing "Crest" with "net weight" would yield the value "5 ounces," assuming this was the size tube being studied. In the most frequently employed variant of the methodology, the specific brand names and information dimensions being used are plainly visible at the outset, while the actual values which result from crossing each brand with each information dimension are concealed.

Subjects are placed into a simulated shopping situation and instructed to purchase one brand from among the set of brands provided in the information display. They are informed that, using the IDB, they are to "shop" for the product as they usually do, taking as much or as little time and acquiring as much or as little information from the matrix as they desire in order to arrive at their purchase decision. Moreover, they can start at and move to any point they wish in the matrix. The instructions also indicate that they need not acquire any information from the matrix and can "purchase" a brand simply on the basis of brand name, if this is how they usually purchase this product. When information values are selected, a process record is formed of the individual's information acquisition behavior which can then be analyzed in terms of the depth (i.e., extent), sequence, and content of the information acquired.

In order to insure that a degree of realistic motivation is present, subjects are told in advance that they should make their decision count since they will receive five "30¢-off" coupons valid only for the specific brand which they select and which can be used in their next five purchases of this product. This cents-off procedure is used in preference to providing subjects with the actual product, because it reduces the likelihood of individuals selecting the highest priced brand even when this is not what they normally would do.

The basic approach is quite flexible and capable of

being modified in a variety of ways in order to address numerous theoretical and applied questions. Two modifications are of particular interest here. First, brand names for established products often serve as information "chunks" (cf. Jacoby, Szybillo, and Busato-Schach, 1974; Simon, 1974), thereby conveying information regarding other attributes of the brand in question. Thus, given the availability of brand name (e.g., "Pledge" furniture polish), subjects often need not behaviorally access certain information values because they already have these in memory (e.g., type of scent: lemon). Accordingly, while the pilot investigation described below makes brand names available for all subjects, some of our earlier and current investigations also use numbers or letters in lieu of actual brand names.

Second, the pilot investigation contains a modification which attempts to approximate reality by introducing a memory load factor in information acquisition. In the basic procedure described above, once information is accessed, it remains visible and available throughout the remainder of the pre-purchase information acquisition decision task. Adding "depth" to each cell on the matrix permits examination of the re-accessing of information values which have not entered long term memory, but have been erased from short term memory. Specifically, this involves constructing the IDB so that each cell in the matrix consists of a three-dimensional pocket. A set of ten identical information value cards is placed into each pocket and arranged so that the backs (i.e., non-information bearing side) of the cards face the subject. A subject interested in acquiring a specific piece of information (e.g., price for Crest) goes to the cell in the matrix containing these values, removes one of the cards, and looks at the information value appearing on the reverse side (e.g., 53¢). Given that each pocket in the IDB contains ten identical cards, the individual may behaviorally re-access information up to ten times, should he so desire. Once a card is acquired and the information looked at, the subject must place it face downward in a collection tray (thus yielding a behavioral process record of information acquisition) before he may acquire another card. Since the subject is always confronted by a display in which all the specific information values are concealed, a memory load factor is imposed.

#### Five Characteristics of the Methodology

Consumer information processing is, by definition, a dynamic phenomenon. Yet virtually all attempts to measure information processing in the consumer context have employed static, cross-section approaches, usually involving verbal reports collected from respondents at a time far removed from the actual information processing and decision-making itself. In contrast, our methodology permits us to capture the depth, content, and particularly the sequence of information acquisition, as it occurs, so that it can later be examined in terms of the dynamic process that it is.

A related characteristic is that it focusses on behavior, not verbal reports. In other words, it focusses on what people actually do, not what they say they do. As many a social scientist will readily testify, behavior and verbal reports of behavior can be (and often are) completely unrelated. In the several studies we have conducted thus far, we generally find only a +.4 (ca.) correlation between the information dimensions that a respondent tells us he refers to and the information values that he actually selects. In our opinion, verbal reports -- particularly those involving recall of long forgotten and/or relatively trivial events (e.g., "Tell me what kinds of information you looked at when you bought your last package of laundry detergent."), and those requesting a description of likely behavior in the

future (e.g., "How likely are you to use nutritional information when you buy frankfurters?") -- are replete with biasing characteristics and are less valid predictors of actual *in-vivo* behavior than are data derived from behavior manifested in the simulated shopping situation just described.

It should also be noted that our methodology is descriptive, not prescriptive. It makes no value judgments; it says nothing about what kind of information should be provided. The procedure simply addresses the issue of just which information is used. In a very real sense, the concern is not so much with the nature of the information provided by the source, as it is with the impact that this information has on the receiver (see Jacoby, 1974, p. 103-106 for an extended discussion of this issue, and Jacoby and Small, 1975, for one concrete application of this approach).

Another aspect of our approach is its attempt to inject a degree of realistic motivation in the decision-making task. While the task is only a simulation, the consequences are real. Something of value to the subject is riding on his decision. If he doesn't like Post Toasties, he had better not select that brand because part of his payment for participating in the study will come in the form of \$1.50 in coupons (i.e., 5 x 30¢-off) good only toward the purchase of that brand. While our subjects are also paid \$3.50 in cash, this payment functions primarily to motivate them to participate in the study. The additional \$1.50 in cents-off coupons is designed to motivate them to take the decision task seriously, that is, to behave realistically while actually engaging in the decision-making task itself.

Finally, the methodology is highly flexible. Two aspects of this flexibility were described above. Several other modifications have been utilized; yet others are currently being developed and evaluated. These will be described in forthcoming papers (e.g., Chestnut and Jacoby, 1975b).

#### A Research Paradigm for the Analysis and Interpretation of Process Data

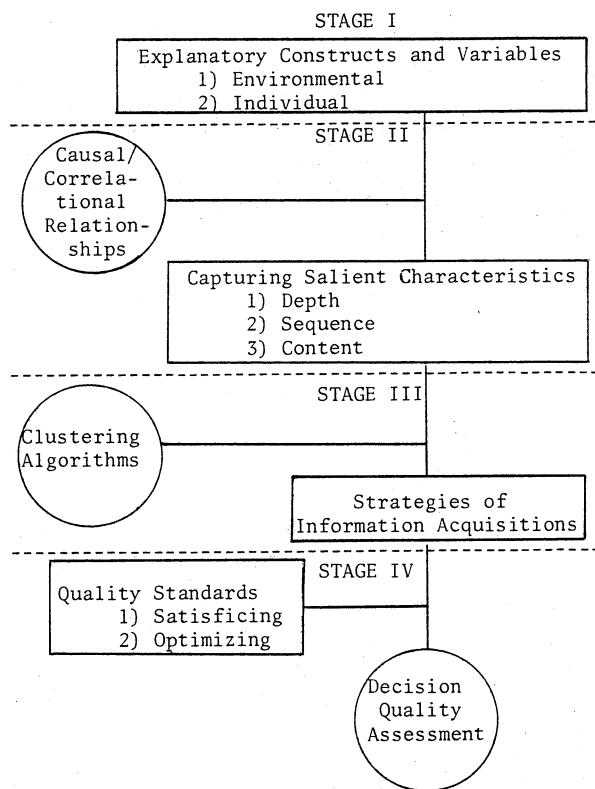
The analysis of process data is appealing from the standpoint of theory. As Simon (1957) notes, only through a complete understanding of the events leading up to a decision can we hope to explain the "basis" for choice. Such understanding, however, is not easily obtained (cf. Lanzetta, 1963). As methodological complexity increases (i.e., in going from static to process measurement), the quantity and nature of the resultant data can begin to exceed the investigator's limited statistical and conceptual models. It becomes necessary to limit research aims and proceed via a logical schema of analysis.

Fortunately, the data analysis problem has, to some extent, already been dealt with in the information processing literature. In the mathematical modeling of decision policies, investigators have typically engaged in a three-step procedure. As reviewed by Dudycha (1970) and Slovic and Lichtenstein (1971), research has attempted: first, to "capture" the decision strategy (i.e., to isolate the salient characteristics of an individual's information processing); then, to "cluster" decision-makers (i.e., grouping individuals according to the homogeneity of their strategies); and, finally, to assess how the clusters or decision strategies relate to some criterion of decision quality (e.g., the "achievement index" of the Brunswik Lens Model). The present paper proposes an extended, four stage version of this paradigm for the study of consumer information acquisition (cf. Figure 1). Each stage focusses on a

different set of variables (represented in rectangles). Attempting to relate one set to another -- making transitions between stages -- highlights important research issues (represented in circles).

FIGURE 1

Four Stage Paradigm for the Analysis and Interpretation of Process Data



Stage I encompasses individual differences (e.g., motivation, uncertainty, perceived risk) and environmental variables (e.g., time pressure, amount of information, distraction) which arouse and direct search behavior. The research issue involved in the transition from Stage I to Stage II is one of establishing causal relationships between specific Stage I variables and the salient characteristics of search strategy. Studies seeking to find such relationships have focussed almost exclusively on environmental variables. In the Newell and Simon tradition (cf. Newell and Simon, 1972), the quality of the "task environment" has been viewed as a major determinant of problem-solving or search activities. Although the environment contains a seemingly infinite number of characteristics, investigators (e.g., Lussier and Olshavsky, 1974, p. 3) have typically maintained that only those characteristics directly relevant to acquisition need be explored. In the realm of consumer behavior, this has led to a recent emphasis upon the environmental phenomena of "information overload" (e.g., Jacoby, Szybillo, and Busato-Schach, 1974; Lussier and Olshavsky, 1974; Payne, 1975, 1976) and "information costs" (e.g., Swan, 1972; Wright, 1974; Winter, 1975; Chestnut and Jacoby, 1975b). Findings almost universally support the impact of such variables upon the nature of consumer information processing. Given the importance of the "task environment" in directing information acquisition (search) behavior, an emerging trend has been to suggest redesigning the consumer's environment so as to optimize the ease and quality of information acquisition (cf. Jacoby, 1974; Bettman, 1975; Russo, 1975).

A factor overlooked in a majority of these studies concerns the potential influence of individual difference variables -- what Newell and Simon would refer to as the problem-solver's "inner-environment." How might motivation, personality, attitudes, or experience affect information acquisition processes? Although investigators have recognized the importance of the individual in terms of "processing" limitations (e.g., Bettman, 1975), few have gone on to study the individual's role in determining the nature of acquisition. Notable exceptions are Swan's (1972) exploration of the effect of product expectations and Weigl's (1975) findings with regard to perceived risk. The pilot investigation described below studied a number of such individual difference variables. Due to space limitations, these are considered in a companion paper (Jacoby, Chestnut, Fisher, and Weigl, 1975).

Stage II involves a consideration of those variables which "capture" (i.e., in some way describe) the salient characteristics of search. Three categories of such variables are depth (i.e., amount of information acquired), sequence (i.e., temporal pattern of information acquired), and content (i.e., specific nature of the information acquired, e.g., price vs. nutrient information).

Given the data obtained from using a "process" methodology, numerous variables and combinative statistics can be generated. Table 1 organizes a variety of such measures in terms of depth, sequence, and content. (N.B. Brief labels are provided in Table 1; more detailed descriptions of selected measures appear in the Appendix.) Some measures are undoubtedly highly correlated with others and could be dispensed with. All are included to emphasize the variety of measurement possibilities. Many have been developed in the course of the present investigation; others are derived from related research independent of our own.

TABLE 1  
Information Acquisition Variables.  
(Stage II)

- I. Depth Variables
  - A. Number of information values chosen
  - B. Number of brands consulted
  - C. Number of information dimensions consulted
  - D. Time taken to arrive at decision
  - E. Mean rated importance of all values chosen
  - F. Mean rank importance of brands consulted
  - G. Mean rated importance of dimensions consulted
  - H. Brand redundancy
  - I. Dimension redundancy
  - J. Brand run index (Bettman)
  - K. Dimension run index (Bettman)
  - L. Proportion of acquired values devoted to brand chosen
  - M. Discrepancy between actual and "optimal" search
  - N. Variance in the number of dimensions considered per brand (Payne, 1975)
- II. Sequence Variables
  - A. Measures based upon acquiring  $\leq$  one value
    1. Immediate brand name choice (no values acquired)
  - B. Measures based upon a sequence of two values
    1. Proportion of type "1" transitions (same brand/same dimension, i.e., immediate re-access)
    2. Proportion of type "2" transitions (same brand-different dimension)
    3. Proportion of type "3" transitions (different brand-same dimension)

4. Proportion of type "4" transitions (different brand-different dimension)
5. Payne's (1976) index (proportion of "2"s minus proportion of "3"s)
6. SAI (Bettman and Jacoby, 1975)
7. SBI (Bettman and Jacoby, 1975)
- C. Measures based upon a sequence  $>$  two values  $<$  full process
  1. Proportion of transition changes
  2. Analysis by segments (e.g., halves, thirds, quarters, etc.)
  3. YX brand selection (Russo and Rosen, 1975)
  4. Proportion of type "5" transitions - type 2; brand previously selected
  5. Proportion of type "6" transitions - type 3; dimension previously selected
  6. Proportion of type "7" transitions - type 4; new brand, old dimension
  7. Proportion of type "8" transitions - type 4; old brand, new dimension
  8. Proportion of type "9" transitions - type 4; old brand, old dimension
  9. Higher order transitions based upon a consideration of 4 or more values
- D. Measures based upon a full process sequence
  1. Transition vector (types 1, 2, 3, & 4)
  2. Expanded transition vector (types 1 through higher order transitions)
  3. Vector of brand choice
  4. Vector of dimension choice
  5. Raw data matrix
- III. Content Variables
  - A. Number of subjects considering a given dimension
  - B. Total number of information values chosen from a given dimension
  - C. Percent of search devoted to a given information dimension
  - D. Behaviorally derived importance weight (Jacoby, Szybillo, Busato-Schach, 1974)
  - E. Percent of intrinsic information in search
  - F. Percent of extrinsic information in search
  - G. Clustering based upon intercorrelation patterns among dimensions acquired

The major research opportunity provided at this juncture suspends interest in explaining search until decision-makers have been satisfactorily "clustered" into homogeneous groups according to their information acquisition strategies. This involves moving to Stage III and raises an important issue. What exactly constitutes a "strategy" of information acquisition? According to our paradigm, strategies reflect some combination of Stage II variables. The problem thus becomes: which combination?

In one of several companion pieces to this paper, Bettman and Jacoby (1975) analyze the pilot data described below by clustering individuals via a sequence definition of strategy. Their primary grouping statistics are the Same Brand Index (SBI) and Same Attribute Index (SAI; see their Figure 1). Russo and Rosen's (1975) YX brand choice pattern (i.e., Bettman and Jacoby's "End Comparison Phase") is entered at selected nodes to achieve a finer gradation of pattern. Once having clustered or defined these strategies, they then proceed backward to Stage I and demonstrate a number of viable relationships (e.g., brand loyal consumers engage in more "Choice by Brand Processing").

Payne (1975) clusters decision-making strategies along somewhat different criteria. Although he too employs the within-brand/within-dimension distinction (i.e.,

inter- vs. intra-dimensional search), a new factor is added: variance in the depth of search per alternative. This provides a slightly different classification rationale and, consequently, new strategies are isolated (e.g., Tversky's "elimination by aspects" and Wright's "conjunctive processing").

Are these the only clustering algorithms of value? Probably not. Table 1 lists a variety of clustering criteria yet to be exploited. Since the actual quality of a decision is more likely to be a function of selected depth and content features, definitions of "strategy" should go beyond mere patterning if the fertility of process data is not to be ignored. Even if sequence were the only salient characteristic of search, indices such as SAI and SBI "capture" a minimal amount of sequence information. Both statistics are defined on an ordering of two consecutively selected information values and would seem inadequate for reflecting the process of any search of substantial duration or complexity. Clearly, the conceptualization of "strategy" needs to be expanded and at least some of the many other criteria for clustering evaluated. Alternative models for defining strategies need to be developed and the interrelationships of these clustering algorithms examined. This should be done with an eye toward "converging operationalism" (cf. Garner, Hake, and Eriksen, 1956). Otherwise, the proliferation of parallel operational definitions will severely impede the accumulation and integration of knowledge. The greatest negative impact would be on any attempt to make progress in terms of Stage IV in our paradigm.

Stage IV consists of variables which act as standards for evaluating the "quality" of information acquisition in terms of the "accuracy" of the resultant decision. As noted in Figure 1, either "optimizing" or "satisficing" standards could be used in such an evaluation (cf. March and Simon, 1958, p. 141). The large scale research program presently underway collects data using both these standards.

The pilot investigation described below has furnished the basis for considerable exploration of the paradigm just described. One thrust has focussed on considering the interrelationships between and among stages. For example, Jacoby, Chestnut, Fisher, and Weigl (1975) examined the relatively neglected issue of how individual difference variables (Stage I) relate to the nature of information search (Stage II). Greater effort, however, has been devoted to exploring appropriate indices for clustering search strategies (Stage III). In developing our process methodologies, it quickly became apparent that the statistics typically employed in consumer research were inadequate for describing and analyzing the data we would be collecting. To exert better leverage on this issue, two independent efforts were mounted to develop suitable process statistics. In one effort, because of his directly relevant earlier work (cf. Bettman, 1970, 1971, 1974a, 1974b), James Bettman was retained as a consultant and given data from the pilot investigation to analyze. The analytical and interpretive work appearing in Bettman and Jacoby (1975) is the output from this effort. Because of space limitations, only one portion of the second effort is described here. Companion pieces (Chestnut and Jacoby, 1975a; Jacoby, Chestnut, Fisher, and Weigl, 1975) provide a more complete accounting.

#### A Pilot Investigation

##### Subjects and Test Product

This study was conducted during the fall of 1974 as a pilot for a larger investigation now in progress. Sub-

jects were 60 Purdue University undergraduates (52 females, 8 males) whose participation met a course requirement. These particular subjects received no other incentives (e.g., cents-off coupons). Cold breakfast cereal, a frequently purchased non-durable, was the test product.

##### Procedure and Apparatus

Subjects, who were interviewed individually, first responded to items assessing product importance, frequency of purchase and consumption, "evoked set," perceived quality differences among brands, perceived healthfulness of cold breakfast cereals, and several other product related indices. Following this, they were introduced to the IDB which was divided into 16 columns (purchase alternatives, i.e., brands) and 35 rows (information dimensions). Purchase alternatives were the 16 brands of cold breakfast cereal having the largest national market shares. Each brand was randomly assigned to a purchase alternative column; each brand name appeared at the top of the column to which it had been assigned.

Information dimensions rows consisted of 35 types of information found on most cold cereal package panels (e.g., price, calories per serving, net weight, etc.). These information dimensions appeared as three alphabetically ordered columns down the center and along the left- and right-hand sides of the IDB. There were thus (35 x 16) 560 cells, each with 10 identical cards specifying an information x alternative crossing. In instances where no information existed regarding a particular information dimension for a given brand, the words "No information" were printed on the reverse side of the cards in that cell.

Subjects were instructed to shop for the product, as described above. The cards each subject deposited in the collection tray provided a behavioral process record (in terms of the depth, sequence, and content) of his pre-purchase information acquisition behavior. Response latency (from the time the first card was touched to the moment a subject announced his purchase decision) was also recorded. Upon completion of the IDB choice task, subjects responded to additional items assessing choice certainty and satisfaction, perceived risk, subjective states experienced during the task (e.g., confusion), recall of information dimensions actually used in information search, rank ordering of the various brands and information dimensions in terms of importance and preferences, and other items relating to the task, the decision making process, and cold breakfast cereal. Finally, a series of open ended debriefing questions sought to determine if demand characteristics were present, and how well subjects comprehended the interviewer's instructions.

##### Analysis

The sequence of acquiring information values can be characterized in a number of ways (cf. Table 1). However, in the relatively few attempts to quantify information acquisition patterns (Bettman and Jacoby, 1975; Payne, 1976), one system of nominal classification has predominated. This approach might be labeled "analysis of transitions" and rests on a consideration of the change in brand and dimension from one acquired information value to the next (i.e., from  $n$  to  $n+1$ ). Focusing upon this relatively simple two item sequence, four possible types of transitions emerge and are distinguished by varying combinations of brand/dimension similarity (see Table 1, IIB, 1-4). Given these basic definitions, a process record of  $n$  values can be expressed as a vector of  $n-1$  transitions with each element in the vector taking on the value 1, 2, 3, or 4, depending upon

the nature of the transition. If vectors vary greatly in length (as in the pilot study) and cannot be adequately compared in terms of overall structure, it becomes necessary to aggregate and describe the pattern of search in terms of the proportion of the vector devoted to each type of transition. Such proportions are then the "observed" probabilities of a given brand/dimension transition in the course of search. Although "analysis of transitions" can be made considerably more complex by accommodating sequences of more than two cards (cf. Table 1, IIc, 5-9), the present investigation restricts itself to the simplest case in order to enhance the clarity of our initial presentation. A more complete analysis is provided in Chestnut and Jacoby (1975a).

Bettman and Jacoby (1975) use a normalized probability measure to cluster subjects into nominal classifications of "strategy." In contrast, the present approach utilizes an interval-like description of search pattern. To achieve this description, five post-hoc models of expected transition probabilities were hypothesized (cf. Table 2). The rationale behind each is as follows. Type 1 transition probabilities were set to zero in all models, since they failed to appear in any great quantity in the observed data (i.e., less than 1% of all transitions). Type 2 and 3 transition probabilities were arbitrarily and systematically varied in order to reflect increasing dominance of either within-brand or within-dimension search (cf. Models I and II, respectively). Type 4 probabilities were increased in two models to indicate search being expanded to include a larger set of brands or dimensions (cf. Models III and IV, respectively). Model V reflects an equal proportion of within-brand (Type 2) and within-dimension (type 3) search. Given such expected probabilities, Chi-square statistics could be computed to describe the degree to which each model "fits" the observed data.

TABLE 2

Definition of Models by Expected Transition Probabilities

Models	Transition Probabilities			
	Type 1	Type 2	Type 3	Type 4
I	.00	.80	.10	.10
II	.00	.10	.80	.10
III	.00	.60	.10	.30
IV	.00	.10	.60	.30
V	.00	.45	.45	.10

Results

Twelve subjects made a purchase decision based only on the available 16 brand names; another subject acquired only one information value. Analysis of the "goodness of fit" between "expected" transition models and "observed" transition probabilities generated five Chi-square statistics for the remaining 47 subjects. Forty-six subjects were categorized under the model for which they had the lowest Chi-square value (cf. Table 3); one subject, whose search consisted of a single type 4 transition (i.e., two values) was not adequately modeled. Mean Chi-square values within each classification were low, indicating the appropriateness of the intuitively chosen probability levels. As search length increased, the model's approximation of the observed probabilities generally improved.

Although not completely similar (for a detailed comparison see Chestnut and Jacoby, 1975a), the clustering results are comparable to Bettman and Jacoby's (1975) dis-

crimination net findings. Combining models I with III and II with IV, we find equal numbers of subjects (17) engaging in either predominantly within-brand (CPB) search or predominantly between-brand (CPA) search. The remaining 12 subjects appear to be just as likely to acquire information through a type 2 transition as they are a type 3 transition (Bettman and Jacoby's CFP strategy).

TABLE 3

Mean Chi-Square Values and Number of Subjects Classified Within Each Model

Models:	I	II	III	IV	V
<b>A. Cumulative</b>					
No. of subjects	N = 9	6	8	11	12
	$\chi^2 = .17$	.19	.12	.10	.18
<b>B. By Search Length</b>					
Light Search, 2 - 6 cards	N = 3	3	2	3	3
	$\chi^2 = .23$	.25	.13	.18	.19
Moderate Search, 7 - 16 cards	N = 3	2	3	5	4
	$\chi^2 = .21$	.18	.18	.08	.19
Heavy Search, 17 - 77 cards	N = 3	1	3	3	5
	$\chi^2 = .06$	.05	.05	.05	.15

Given the rather clear-cut search patterns that emerge, it seems reasonable to expect causal factors to be operating. Jacoby, Chestnut, Fisher, and Weigl (1975) present data bearing on a number of potentially influential individual difference variables. Two, in particular, seem promising: brand loyalty and consumption frequency. Table 4 presents a path analysis (cf. Van de Geer, 1974) of the impact of these variables upon search patterning. Five multiple regressions were computed utilizing the Chi-square index of model fit as the criterion, with consumption frequency (operationally defined by a 7-interval scale ranging from "more than once a day" to "less often than once every two weeks") and brand loyalty (defined in terms of the percent-of-purchase devoted to the leading brand) as the predictors. Since the correlation between predictors was -.06 (n=48), they were assumed to act independently. Path coefficients could therefore be interpreted in terms of a causal model. Because subjects in the "light search" category were inadequately modeled by the Chi-square statistic (cf. Table 3) and could obscure the results, searches of less than seven values were excluded from the analysis. This reduced sample size to 32.

TABLE 4

Multiple Correlations and Path Coefficients<sup>1</sup> for a Prediction of Model Fit

Model	$r_{mult.}$	(n=32) sig.	path coefficients	
			heavy consumpt.	brand loyalty
I	.51	.01	.51	.02
II	.42	.06	-.27	-.29
III	.45	.03	.45	.00
IV	.43	.05	-.29	-.28
V	.47	.02	.40	-.30

<sup>1</sup>Positive path coefficients reflect an increase in model fit.

The resultant multiple correlations were moderate, ranging in significance from .01 to .06. Path coefficients indicated three types of predictor influence. Within-brand (i.e., type 2 search, models I and III) was best predicted by consumption frequency; specifically, the more often the product was consumed, the greater the

tendency to engage in type 2 search. The remaining models were determined by a combination of consumption frequency and brand loyalty. Between-brand comparison (i.e., type 3 search, models II and IV) was more likely with non-loyal consumers having a low frequency of consumption. Complex strategies (model V), on the other hand, were employed by nonloyal consumers having a high frequency of consumption. Although these relationships are significant, the unexplained variance attributable to latent constructs not made explicit in the path analysis should be considered. Approximately 70 to 80 percent of the variance in model fit remains unexplained. However, despite sample limitations and self-report scaling of the predictor variables, the results seem promising and warrant further examination (see Jacoby, Chestnut, Fisher, and Weigl, 1975).

#### Discussion and Conclusions

Quite clearly, there appear to be distinct differences in the types of information acquisition strategies consumers employ. Different consumers use different information acquisition strategies. Whether these same patterns emerge with other kinds of products and consumers, as well as the causal factors underlying these patterns, is currently being examined. Data regarding the relationship of these patterns to criteria of decision quality (cf. Stage IV in our paradigm) are also being collected.

Another apparently inviolate generalization emerging from this and other investigations (e.g., Jacoby, Szybillo, and Busato-Schach, 1974; Olson and Jacoby, 1972; Payne, 1975) is that consumers typically acquire only a small proportion of the available information. Utilizing a 35 x 16 matrix display of package information relevant to breakfast cereal, we find a median search length of less than seven information values. Of the 560 values available, half the sample found it necessary to acquire less than 2% of the information available in the display prior to making their purchase decision. Even if we correct for the fact that brand names were exposed (i.e., could be treated as an acquired information dimension) and that some of the dimensions were more relevant to pre-usage and pre-disposition decision making (thereby reducing the matrix to a conservative estimate of 16 pre-purchase dimensions such as price, net weight, etc.), a selective exposure of less than 9% can be estimated. Given that the methodology employed, if anything, should have increased search length (i.e., through potential demand characteristics and the minimization of acquisition costs relative to such costs in an actual supermarket setting), this degree of selectivity is striking. It is of course possible that using brand name as an information chunk (cf. Jacoby, Szybillo and Busato-Schach, 1974; Bettman and Jacoby, 1975) may be partially responsible for some of this selectivity in this investigation, and ongoing research is exploring this question.

When consumers do acquire package information, however, an interesting relationship appears through an analysis of Stage I individual difference variables and their effect upon selected sequence statistics at Stage II. Specifically, consumption frequency and brand loyalty are seen to determine the relative proportions of type 2 and type 3 transitions. One explanation for this relationship might be framed in terms of a "stimulus recoding" interpretation of the information acquisition process. If we assume that the type of stimulus recoding employed in the search task is a function of the nature of the information stored in long-term memory (cf. Massaro, 1975, p. 249), it would seem only reasonable to expect a relationship between selected purchasing characteristics (i.e., as correlates of knowledge) and the pattern of information acquisition. That is,

a person who consumes the product quite frequently is more likely to have stored information about the properties of the product and, thus, employ a within-brand recoding which would input values of information in terms of well-known properties related to a single brand. In contrast, a person who is non-loyal could be expected to be more aware of the many different brands present in the product class (cf. Jacoby and Hillemeier, as cited on p. 206 of Jacoby, 1975) and, thus, resort to a within-dimension recoding which would input values of information in terms of well-known brands that emerge successful from a consideration of a single dimension. Much as in the paired-associate analysis of mnemonic techniques (cf. Paivio, 1971), the subject is utilizing a conceptual peg of personal salience on which to hang (i.e., remember) a specific item of information. Such an interpretation fits the empirical findings of the path analysis and would seem to represent a worthwhile topic for further research.

With regard to our analytic method, the Chi-square modeling approach to clustering subjects is but one of many ways to examine information acquisition patterns. Bettman and Jacoby's decision net approach is another. Payne's reliance on the variance in the depth of search per alternative is yet another. Probably the most important distinction between these approaches is that the Chi-square method generates interval-like data, thereby providing a more satisfying means for relating transition probabilities to other variables. Yet other clustering approaches are described and employed in Chestnut and Jacoby (1975a).

Regardless of which clustering algorithm is used, arbitrary assumptions are usually involved (cf. Cormack, 1971). Although these assumptions may describe the data at hand, they need not reflect the true nature of reality. The stability of derived clusters and validity of inferences drawn from their analysis need to be further established. As indicated, ongoing investigations are focussing on cross-validation and on the extension of algorithms to include depth and content criteria. If the same relatively distinct clusters emerge, empirical description can give way to a more theoretical analysis since, as Hartigan (1975) notes, "clear-cut and compelling clusters...require an explanation of their existence and so promote the development of theories..." (p. 7).

A limitation common to both the Chi-square modeling and decision net approaches is that they reduce the dynamic information acquisition process to a static representation. Now that consumer research is beginning to develop methodologies which capture the dynamic character of the information acquisition process, a set of analytical statistics which preserve this character is needed. One such approach is described in Chestnut and Jacoby (1975a).

A final point should be noted in regard to the place of verbal protocols in our methodology. As in many of the quantitative modeling investigations (cf. Wright, 1972), our subjects are questioned about their search strategy only after the completion of their decision. This is in marked contrast to related work by others (e.g., Bettman, 1971; Payne, 1975; and Lussier and Olshavsky, 1974) in which verbal protocols of search are collected concurrent with behavioral simulation. The reasons we adopted our post hoc procedure are threefold. First, our intent has been to develop a minimally disruptive and reactive approach to measuring the actual acquisition process. By not forcing the subject to verbalize or provide rationale behind his acquisition behavior, we lessen the possibility that his behavior and protocols will reflect biasing demand characteristics and leave him free to interact with the information display in a manner more akin to actual package informa-

tion search. Second, because of the unwieldy volume of data that is generated per subject, methodologies employing verbal protocol tracing procedures have typically been forced to consider very small numbers of subjects (e.g., n=6 and 12 in Payne, 1975; n=2 in Bettman, 1970). In contrast, the present study utilized 60 subjects; our ongoing investigation involves a sample size in excess of 750. Finally, it is an open question as to exactly how much of what a subject verbalizes about the purchase of a relatively minor, frequently purchased non-durable such as breakfast cereal is actually clear and meaningful to the consumer himself. In contrast, decision criteria and information acquisition and processing strategies with respect to infrequently purchased and relatively important purchase decisions (e.g., purchase of a home, renting an apartment) are likely to be more cognitively focussed and clear to the decision maker. Verbal protocols should be much more meaningful and revealing in these instances. It should be obvious, however, that studies are needed which comparatively test the impact of collecting verbal reports concurrently vs. post hoc.

Above and beyond the few concrete "findings" presented, we view this paper as making several contributions. In particular, we point to the description of our process methodology, the presentation of our four stage research paradigm, and the development of the Chi-square modeling approach. Hopefully, another contribution will be the heuristic influence this paper will have on future research regarding consumer information processing.

## Appendix

### Definitions of Selected Stage II Variables

- I.E. Mean rated importance of all values chosen: Given numerical importance ratings (typically obtained via a standard Likert scaling) for all dimensions and brands, each information value choice can be represented by the product of the two ratings. An average of these products over some duration of search is then computed.
- I.F. Mean rank importance of brands consulted: If brands are rank ordered, an average rank of brands consulted in search can be computed.
- I.H. Brand redundancy: The number of times a subject returns to a brand previously examined.
- I.J. Brand Run index: Defining a "brand run" as any sequence of two or more values on the same brand, this index can be defined as the number of values chosen minus the number of brand runs divided by the number of values chosen minus the number of brands chosen.
- I.L. Proportion of acquired values devoted to brand chosen: Given that the subject is involved in a decision-making task resulting in the purchase of a single brand, the proportion of search devoted to the brand eventually purchased can be calculated.
- I.M. Discrepancy between actual and "optimal" search: Given ratings of the number of brands thought to be "acceptable" and the number of dimensions thought to be "important for purchase", an estimate of "optimal" search is simply the product of the two ratings. The above index calculates the discrepancy by subtracting the optimal search estimate from the actual number of information values chosen.
- II.C.1. Proportion of transition changes: The frequency of change in type of transition divided by the number of transitions in the transition vector.
- II.C.2. Analysis by segments: Search length can be standardized by dividing the transition vector into a unit length (halves, thirds, quarters, etc.). Given this partitioning, a predominant transition per unit length can be designated and an analysis of the result-

ing equal-length vectors attempted.

II.C.3. XXY brand selection: This is simply the frequency of any XXY pattern in the vector of brand choice.

II.D.3. Vector of brand choice: This is a  $1 \times q$  vector where the column-order is the number of values chosen and each element is a digit representing a specific brand.

II.D.4. Vector of dimension choice: This is a  $1 \times q$  vector where the column-order is again the number of values chosen and each element is a digit representing a specific dimension of information.

II.D.5. Raw data matrix: This is a  $2 \times q$  matrix where the column-order is the number of information values chosen and the two row vectors are the vectors of brand and dimension choice.

III.D. Behaviorally derived importance weight: The sum of the position ranks for choices relating to a given dimension divided by the number of times the dimension was chosen. This can then be standardized across different search lengths by dividing by the number of information values chosen. For example, in a 10 value search, the 1st and 5th value might relate to price. An index of price information importance would then be calculated as follows:  $[(1+5) \div 2] \div 10 = .3$

III.E. Percent of intrinsic information in search: Intrinsic information is defined as any information dimension for which a change in value would result in a direct change in the product's physical composition (cf. Szybillo and Jacoby, 1974). Calories, therefore, would be classified as intrinsic whereas price might be thought of as extrinsic information. Given a classification of each dimension in terms of these two categories, the relative proportion of search devoted to each category can be calculated.

III.G. Clustering based upon intercorrelation patterns among dimensions acquired: Given a  $p \times q$  matrix of subjects by dimensions, where individual entries represent the number of times a given subject consulted a given information dimension, a Q-factor analysis can be used to cluster subjects by general type of information chosen.

## References

- Carol A. Berning and Jacob Jacoby, "Patterns of Information Acquisition in New Product Purchases," Journal of Consumer Research, 1(September, 1974), 18-22.
- James R. Bettman, "Information Processing Models of Consumer Behavior," Journal of Marketing Research, 7(August, 1970), 370-76.
- James R. Bettman, "The Structure of Consumer Choice Processes," Journal of Marketing Research, 8(November, 1971), 465-71.
- James R. Bettman, "Toward a Statistics for Consumer Decision Net Models," Journal of Consumer Research, 1(1974), 71-78.
- James R. Bettman, "Decision Net Models of Buyer Information Processing and Choice: Findings, Problems, and Prospects," in G. D. Hughes and M. L. Ray, eds., Buyer/Consumer Information Processing (Chapel Hill, North Carolina: University of North Carolina Press, 1974, 59-74).
- James R. Bettman, "Issues in Designing Consumer Information Environments," Working Paper Series, Center for Marketing Studies, No. 32 (UCLA, July, 1975).
- James R. Bettman and Jacob Jacoby, "Patterns of Processing in Consumer Information Acquisition," in B.B. Anderson, ed., Advances in Consumer Research, Vol. 3 (Proceedings



- of the Sixth Annual Conference of the Association for Consumer Research, Cincinnati, Ohio, October 30-November 2, 1975).
- Robert W. Chestnut and Jacob Jacoby, "Analytical Techniques for Examining Process Data," Purdue Papers in Consumer Psychology, No. 154, 1975a.
- Robert W. Chestnut and Jacob Jacoby, "The Impact of 'Time Costs' upon the Acquisition of Package Information," Purdue Papers in Consumer Psychology, No. 155, 1975b.
- R. M. Cormack, "A Review of Classification," Journal of the Royal Statistical Society, A134 (1971), 321-67.
- Arthur L. Dudycha, "A Monte Carlo Evaluation of JAN: A Technique for Capturing and Clustering Raters Policies," Organizational Behavior and Human Performance, 5(5, 1970), 501-516.
- Howard L. Fromkin and Siegfried Streufert, "Laboratory Experimentation," in M.D. Dunnette, ed., The Handbook of Industrial and Organizational Psychology (Chicago: Rand-McNally, 1976).
- Wendell R. Garner, Harold W. Hake, and Charles W. Ericksen, "Operationism and the Concept of Perception," Psychological Review, 65(May, 1956), 149-59.
- John A. Hartigan, Clustering Algorithms (New York: Wiley, 1975).
- Jacob Jacoby, "Consumer Reaction to Information Displays: Packaging and Advertising," in S.F. Divita, ed., Advertising and the Public Interest (Chicago: American Marketing Association, 1974, 101-118).
- Jacob Jacoby, "Perspectives on a Consumer Information Processing Research Program," Communication Research, 2(3, 1975), 203-15.
- Jacob Jacoby, Carol Kohn Berning, and Thomas F. Dietvorst, "What about Disposition?," Purdue Papers in Consumer Psychology, No. 152, 1975.
- Jacob Jacoby, Robert W. Chestnut, William Fisher, and Karl C. Weigl, "Individual Difference Factors in Consumer Information Acquisition," Purdue Papers in Consumer Psychology, No. 153, 1975.
- Jacob Jacoby and Constance B. Small, "The FDA's Approach to Defining Misleading Advertising," Journal of Marketing, 39(October, 1975), 65-68.
- Jacob Jacoby, George J. Szybillo, and Jacqueline Busato-Schach, "Information Acquisition Behavior in Brand Choice Situations," Purdue Papers in Consumer Psychology, No. 140, 1974. Provisionally accepted for publication in the Journal of Consumer Research.
- John T. Lanzetta, "Information Acquisition in Decision Making," in O.J. Harvey, ed., Motivation and Social Interaction: Cognitive Determinants (New York: Ronald Press, 1963, 239-65).
- Denis A Lussier and Richard W. Olshavsky, "An Information Processing Approach to Individual Brand Choice Behavior," paper presented at the ORSA/TIMS Joint National Meeting, San Juan, Puerto Rico (October, 1974).
- James G. March and Herbert A. Simon, Organizations (New York: Wiley, 1958).
- Dominic W. Massaro, Experimental Psychology and Information Processing (Chicago: Rand McNally, 1975).
- Allen Newell and Herbert A. Simon, Human Problem Solving (Englewood Cliffs, N.J.: Prentice-Hall, 1973).
- Jerry C. Olson and Jacob Jacoby, "Cue Utilization in the Quality Perception Process," in M. Venkatesan, ed., Proceedings, Third Annual Conference of the Association for Consumer Research (1972), 167-79.
- Allan Paivio, Imagery and Verbal Processes (New York: Holt, Rinehart & Winston, 1971).
- John W. Payne, "Heuristic Search Processes in Decision Making," in B.B. Anderson, ed., Advances in Consumer Research, Vol. 3 (Proceedings of the Sixth Annual Conference of the Association for Consumer Research, Cincinnati, Ohio, October 30-November 2, 1975).
- John W. Payne, "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," Organizational Behavior and Human Performance (1976), forthcoming.
- J. Edward Russo, "An Information Processing Analysis of Point-of-Purchase Decisions," Proceedings, American Marketing Association Fall Conference (1975), forthcoming.
- J. Edward Russo and Larry D. Rosen, "An Eye Fixation Analysis of Multi-Alternative Choice," Memory and Cognition, 3(May, 1975), 267-76.
- Herbert A. Simon, Models of Man (New York: Wiley, 1957).
- Herbert A. Simon, "How Big is a Chunk?," Science, 183 (February 8, 1974), 482-88.
- Paul Slovic and Sara Lichtenstein, "Comparison of Bayesian and Regression Approaches to the Study of Information Processing Judgment," Organizational Behavior and Human Performance, 6(6, 1971), 649-744.
- John E. Swan, "Search Behavior Related to Expectations Concerning Brand Performance," Journal of Applied Psychology, 56(August, 1972) 332-35.
- George J. Szybillo and Jacob Jacoby, "Intrinsic vs. Extrinsic Cues as Determinants of Perceived Product Quality," Journal of Applied Psychology, 59(1, 1974), 74-78.
- John P. Van de Geer, Introduction to Multivariate Analysis for the Social Sciences (San Francisco: W. H. Freeman and Co., 1971).
- Karl C. Weigl, "Perceived Risk and Information Search in a Gift Buying Situation," unpublished M.S. thesis, (Purdue University, 1975).
- William L. Wilkie, "Assessment of Consumer Information Processing Research in Relation to Public Policy Needs," draft report for the Office of Exploratory Research and Problem Assessment, National Science Foundation Grant GI-42037), June, 1974
- William L. Wilkie and David M. Gardner, "The Role of Marketing Research in Public Policy Decision Making," Journal of Marketing, 38(17, 1974), 38-47.
- Frederick W. Winter, "Laboratory Measurement of Response to Consumer Information," Faculty Working Papers, School of Commerce, No. 227(University of Illinois, Jan.10, 1975).
- Peter L. Wright, "Consumer Judgement Strategies: Beyond the Compensatory Assumption," in M. Venkatesan, ed., Proceedings, Third Annual Conference of the Association for Consumer Research, 1972, 316-24.
- Peter L. Wright, "The Harassed Decision Maker: Time Pressures, Distractions, and the Use of Evidence," Journal of Applied Psychology, 59(5, 1974), 555-61.

PATTERNS OF PROCESSING IN CONSUMER  
INFORMATION ACQUISITION

James R. Bettman, University of California, Los Angeles  
Jacob Jacoby, Purdue University<sup>1</sup>

Abstract

Patterns of information acquisition from a brands x information dimension matrix were studied for sixty subjects. Three major patterns were found for those subjects seeking information: processing by attributes, by brands, and a hybrid form, feedback processing. Classification procedures for these patterns were developed. Relating these patterns to other variables was explored.

Despite recent interest in consumer information processing (Hughes and Ray, 1974), perceptual and memory phenomena have not been handled well in most models. These include: perceptual encoding, information acquisition, storage, and retrieval. These areas have been virtually ignored in consumer information processing research. For example, consumer decision net models (Bettman, 1970) use an extremely simplistic memory structure. Memory is modeled as a fixed, given data array. Encoding is ignored, as is acquisition. The data are there to be utilized, represented as elements in a matrix. Retrieval is assumed to be automatic and error-free. As Calder (1975) points out, how information is stored in memory can effect the type of processing rule used. Even if the rules used in making choices are of primary concern, perceptual and memory issues must be examined. This paper focuses on patterns of information acquisition, only briefly considering storage and retrieval.

Study of acquisition patterns is necessary for a theory of consumer information processing. It also has pragmatic value for study of public policy issues related to provision of information to consumers. Tasks such as reading labels on packages in a supermarket or looking at a table in a Consumer Reports article are tasks of information acquisition. As Russo, Krieser, and Miyashita (1975) show, how information is presented in such a task can greatly affect whether the information is used. Knowing how consumers acquire information from data displays can aid in developing more effective displays (Russo, 1975).

Some recent research in both cognitive psychology and consumer decision making has begun to attack issues of information acquisition (Berning and Jacoby, 1974; Jacoby, Szybillo, and Busato-Schach, 1974; Lussier and Olshavsky, 1974; Payne, 1976; Russo and Rosen, 1975; Simon and Barenfeld, 1969; Svenson, 1974.)

A common focus is on the order and patterns of information acquisition during the choice process. Some stable findings emerge, although often presented without formal analysis. First, there are individual differences in order of information search and acquisition. Some subjects proceed by examining one brand at a time. That is, they choose a brand and gather information on several attributes of that brand. Then they choose a second brand and gather information on several attri-

butes (not necessarily the same as those for the first brand) and so on. This strategy may be called Choice by Processing Brands (CPB).<sup>2</sup> A second group acquires information by choosing an attribute, determining values for each of several brands on that attribute, choosing a second attribute and determining values for several brands and so on. This may be called a Choice by Processing Attributes (CPA) strategy. Finally, some subjects engage in hybrid strategies. These strategies seem to be adaptive, governed by immediate feedback from the information found during the search. Alternating short sequences of brand and attribute processing are often observed (e.g., a subject may process brands 2, 3, and 4 for attribute 1, become interested in brand 4 and next process attributes 2 and 5 for it, then examine attribute 5 for brands 2 and 3, and so on.) This may be termed a Choice by Feedback Processing (CFP) strategy.

A second set of findings is that some subjects use the basic strategies above in a uniform manner; they do not change their strategy during the course of the search. Other subjects use phased strategies (Wright, 1974), where the type of processing varies. A common type of phased strategy is one where the first phase is information input, and the second phase is a set of paired comparisons among specific alternatives (Svenson, 1974; Lussier and Olshavsky, 1974; Russo and Rosen, 1975; Payne, 1976).

This study examined information acquisition in a choice of breakfast cereal brands. Two methodological issues must be dealt with: first, the task must be designed so that the data on the sequence of acquisition are available; second, measures must be developed that allow classification of patterns of acquisition. The efforts described here represent advances over previous work.

Method

The details of the procedures employed in this study are provided in Jacoby, Chestnut, Weigl, and Fisher (1976). Suffice it to say that it involves the presentation of information via a brand x information dimension matrix display and the sequential acquisition of information from this array by 60 subjects involved in a simulated breakfast cereal purchasing situation. By examining the pile of information cards acquired by each subject, a record of the information acquisition process as a sequence of brand-attribute pairs could be developed.

This task is essentially the same as that used by Payne (1976) in his apartment choice study. Payne also collected verbal protocols from his subjects. Russo and Rosen (1975) used eye movements, and Lussier and Olshavsky (1974) and Svenson (1974) used only verbal protocols. The present methodology yields a behavioral record of the complete external information search (internal retrieval from long term memory was, of course, not monitored), as

<sup>2</sup>The terms 'Choice' and 'Processing' are used, although acquisition is the primary focus, because subjects are asked to choose an alternative after their information search. Hence, acquisition and choice processes are confounded. This is discussed further below.

<sup>1</sup>Preparation of this report was facilitated, in part, by a grant from the National Science Foundation (GS-43687) to the junior author. A more complete working paper version of this article is available from either author.

does the eye movement data. Protocol data has value in potentially illuminating some aspects of internal search and rationale for the external search, although the record of the external search obtained may be incomplete. The combination of an information display board and protocols would obtain fairly complete information, but be cumbersome to use with many subjects.

The main differences between the method of the present study and those used in past studies are the larger sample size (60 compared to 12 [Russo and Rosen, 1975]; 6 and 12 [Payne, 1976]; 6 [Svenson, 1974]; and 27 [Lussier and Olshavsky, 1974]); and the use of actual brands as choice objects and actual information values. The previous studies have used hypothetical alternatives.

#### Measuring Acquisition Patterns

The major datum available for classifying information acquisition patterns is the sequence of information cards (for each subject who took cards), which yields a sequence of brand-attribute pairs. The only previous attempt at developing an index to measure the structure of this input acquisition sequence was by Payne (1976), based upon examining adjacent pairs of brands and of attributes in the sequence. If  $n$  cards are chosen in the task, there are  $n-1$  pairs of brands and of attributes. Each pair can consist of two identical brands (attributes) or two different ones. Let  $SB$  equal the number of identical brand pairs divided by  $n-1$ , and let  $SA$  be defined similarly for attributes. Then Payne's index is defined by  $(SB-SA)/(SB+SA)$ . For a subject following a pure Processing Brands (CPB) strategy, this index would be +1. For a pure Processing Attributes (CPA) strategy, it would be -1. Payne also developed an index of shifts in the sequence. A shift occurs whenever neither a brand nor an attribute is common for two adjacent cards in the sequence. The proportion of shifts is approximately given by  $1-SA-SB$ .

Payne's index for classifying subjects does not take into account that values of the  $SA$  and  $SB$  measures have different ranges for different numbers of cards chosen and total number of attributes and brands considered. With 10 cards chosen and 4 brands in total considered, the maximum value possible for  $SB$  is  $(10-4)/9$  or .667, for example. In this study, normalization was performed by dividing  $SB$  and  $SA$  by the maximum possible for each subject, given the number of cards chosen and numbers of brands and attributes considered. These normalized measures are denoted by  $SBI$  (Same Brand Index) and  $SAI$  (Same Attribute Index). Both range from 0 to 1. This focuses on a structural criterion, rather than using amount or specific content of the information chosen.

These indices were used to determine the basic form of strategy used. In addition, however, the phased nature of the strategies was examined. The major phased strategy was where a paired comparison between brands occurred near the end of the sequence. The rule used to identify paired comparisons was that used by Russo and Rosen (1975) (their weak criterion): that brand alternations of the form  $X-Y-X$  appear in the sequence. The rule to determine whether a phased strategy (end comparison) was being used was whether or not a subject's sequence contained at least one paired comparison in the last six entries.

The classification procedure was exploratory. One of the authors (JB) examined each subject's sequence in detail. An attempt was then made to model the pattern classification rule used in making his judgments. For this particular author, this naturally resulted in a discrimination net, given in Figure 1. Although clustering algorithms based on the indices discussed above and

other possible indices could clearly be applied, it was felt that at this stage of the research using more flexible human pattern recognition capabilities was desirable.

This classification procedure is based mainly upon the indices discussed above. First, subjects who chose no cards are classified as Choice by Brand Name (CBN). Second, subjects who used only one attribute or one brand are singled out as special cases of the CPA or CPB strategies, CPA-1 and CPB-1. Then, if the Same Brand Index (SBI) is high and the Same Attribute Index (SAI) is low, Choice by Processing Brands (CPB) is selected as the appropriate pattern. If SAI is high and SBI low, then Choice by Processing Attributes (CPA) is the classification. If both SAI and SBI are moderately high, Choice by Feedback Processing (CFP) is indicated. Note that for CFP to be as it was characterized above (alternating short same brand and attribute sequences), the number of shifts should not be a great deal higher than for CPA or CPB. Phased strategies with an end comparison (EC) are checked for in each of these instances.

This classification scheme is one of the only "formal" procedures presented in the literature. Most studies simply report what subjects appear to do without measuring processing by brands versus processing by attributes, described above. In his second experiment, Payne used his index to classify subjects, but he only used two groups, comparable to the attribute processing and brand processing groups of this study. This does not recognize feedback processing as a separate strategy, but confounds it with the other two. This is unfortunate, since the feedback processing group seems to have interesting and distinct properties, as shown below. The procedure used in the present study adds further precision in classifying subjects.

#### Results

The results of applying the discrimination net to the data from the sixty subjects are as follows, using the abbreviations in the key for Figure 1: CBN, 12; CPB, 11; CPB-1, 4; CPB-EC, 3; CPA, 6; CPA-1, 5; CPA-EC, 11; CFP, 6; CFP-EC, 1; and Other, 1. There is relatively high use of Choice by Brand Name. Approximately equal numbers of CPB and CPA strategies were utilized, with more phased strategies for CPA users. Use of the CFP strategies is more limited.<sup>3</sup> These classification results present a more formal view of individual differences in acquisition processing than past research, and provide insights into the empirical extent of different types of processing lacking in the previous studies. However, since only one product class was studied, the results may be dependent on this specific class.

In an attempt to characterize each strategy type more closely, three main types were examined for relationships with other variables. The various versions of CPA, CPB and CFP were collapsed together, to form groups of 18(CPB), 22(CPA), and 7(CFP). The analyses reported below were carried out for these 47 subjects. Choice by Brand Name subjects were not included in the analyses because they did not have most of the variables utilized available. Formal analyses of relationships between

<sup>3</sup> The depiction of Choice by Feedback Processing utilized above was supported by the shift index. Values of the index were .154 for the 18 CPB subjects; .194 for the 22 CPA subjects, and .197 for the 7 CFP subjects ( $p < .66$  in a one way analysis of variance). The average values of Payne's classification index for the three groups were .780 (CPB), -.772 (CPA), and .034 (CFP). The nearly zero value for the feedback processing group shows the confounding effects of Payne's classification scheme.

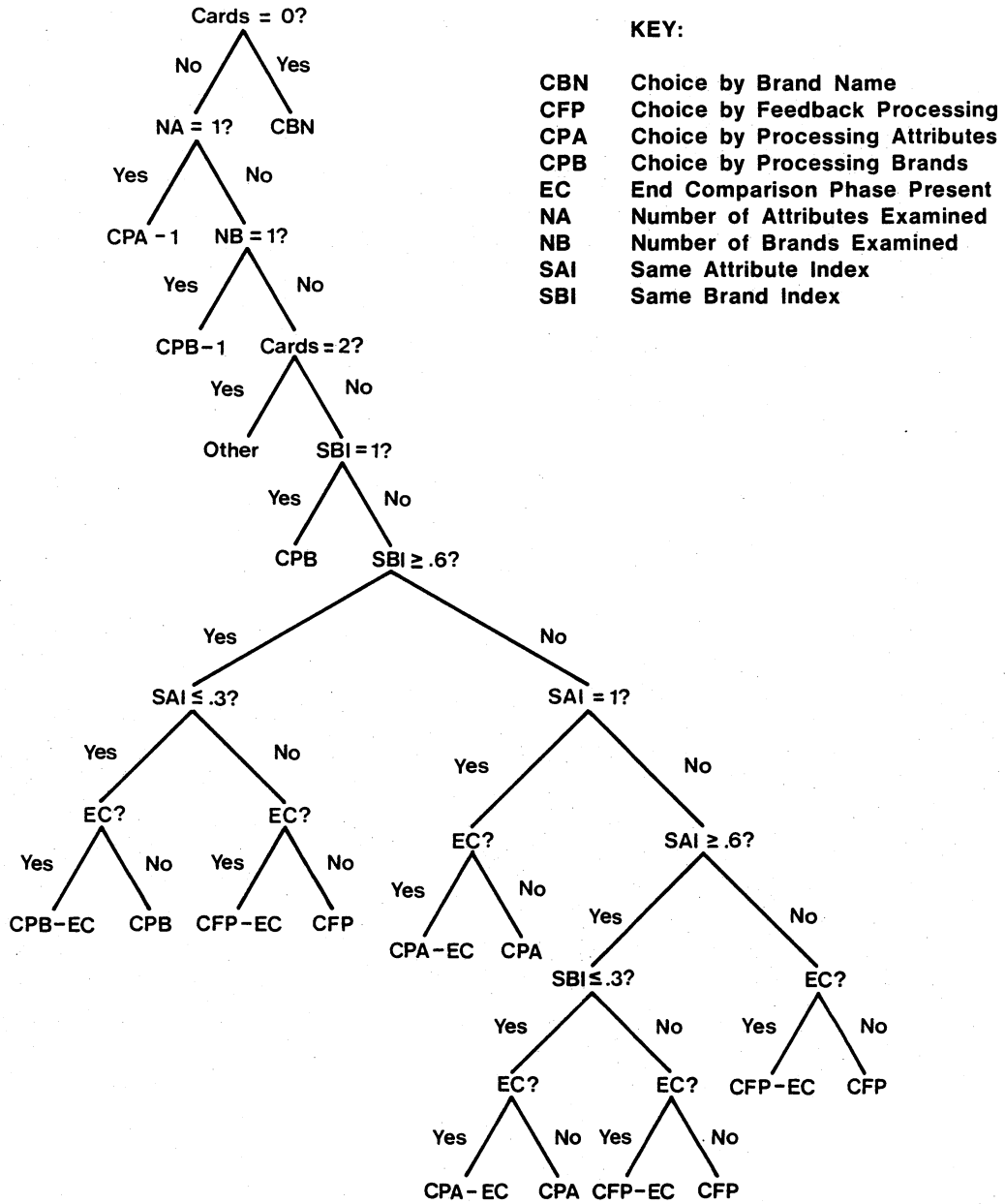


Figure 1. DISCRIMINATION NET FOR ACQUISITION STRATEGIES

process type and other measures have not been reported in previous studies.

One way analyses of variance were performed for each of several variables using the three groups as factor levels. The mean values for each group and  $\alpha$ -level probabilities for these ANOVAS on each variable follow: Number of Cards Chosen-CPB, 12.39; CPA, 13.09; CFP, 22.43 ( $p < .19$ ); Amount of Time Taken (Seconds)-CPB, 201.89; CPA, 178.05; CFP, 313.86 ( $p < .15$ ); Number of Brands Examined-CPB, 2.89; CPA, 4.86, CFP, 5.71 ( $p < .05$ ); Number of Attributes Examined-CPB, 6.06; CPA, 3.64; CFP, 6.29 ( $p < .04$ ); Perceived Quality Differences (1=very large, 5=hardly any)-CPB, 3.50; CPA, 3.41, CFP, 2.57 ( $p < .18$ ). These results show that subjects using a Choice by Processing Brands strategy tend to examine fewer brands; those using a Choice by Processing Attributes strategy tend to use fewer attributes. These subjects tend to concentrate in depth on a few of the aspects they are using to organize their search (i.e., brands or attributes). Choice by Feedback Processing subjects may be using a more detailed acquisition process. They tend to use more time, take more cards, look at more brands and attributes, and see greater quality differences among brands in the product class. These findings are of course preliminary, since the data set is small, the significance levels only suggestive, and they are selected from a larger set of findings, but they provide hypotheses for future research. They are also supported by a contingency table relating processing strategy and a brand loyalty index (one if one brand bought more than 50% of the time, two otherwise). For CPB, 14 of 18 are brand loyal, for CPA 10 of 21, and for CFP, 1 of 7. This table yields a  $X^2$  significant at  $p < .02$  (although expected cell sizes for the CFP subjects are small).

A consistent pattern seems to emerge: subjects using CFP strategies seem to engage in the most explicit processing, CPA subjects in an intermediate amount, and CPB subjects in a lesser amount. The CPB strategy does not explicitly include across brand comparisons (although internal processing may be high), the CPA strategy includes such comparisons for each attribute, and the CFP strategy seems to be a more highly external and evaluatively guided search process. To more fully understand the actual effort allocated in each case, however, evidence on internal as well as external processing would be necessary.

#### Implications

##### Properties of the Task

Some biases inherent in the task used must be discussed, since they may affect the results. First, since there is no limitation on the number of cards a subject can choose, there is no necessary incentive for the subject to "optimize" his search process, or organize his or her search. There is no explicit time pressure involved, and subjects may in fact be resigned to spending time 'in the lab'. Hence, search may be less directed than in a normal shopping environment. A more serious problem may be the availability of brand name to the subjects. This availability does coincide with the real world; however, most studies of cognitive strategies attempt to create situations where extra-experimental knowledge and differences in long-term memory functioning are not factors in performance. In this task, however, differential familiarity with brand names could affect the results. Also, "check up" strategies may be employed, where subjects may choose cards to verify information they already have stored in long-term memory (this may be true of actual store decisions also). These factors could bias the form of the process toward

Choice by Brand Name or Choice by Processing Brands.

Another potentially biasing factor, mentioned above, is that the sequence of cards chosen confounds acquisition, evaluation, and choice processes, since subjects were required to choose a brand. However, this confounding seems to be inevitably characteristic of actual consumer decision making and experiments on choice, although Russo and Rosen (1975) present evidence that some of the effects can be unconfounded. In the present study, there may be a good deal of evaluative processing done internally, without ever being reflected in the card sequence explicitly.<sup>4</sup>

Finally, the experimental setup makes it equally easy, in terms of the effort needed to acquire the information, to process by brands or by attributes or to use a feedback strategy. However, this property is not true of the real world, except for perhaps tasks such as reading a table in Consumer Reports. For a choice of brands from a supermarket shelf, information is organized by brand more than by attribute. Thus perceptual encoding by brand is facilitated; this may in fact hinder evaluative brand comparison processes.

##### Implications for Consumer Information Processing Models

The results show that over half of the subjects process by attributes (CPA) or by a combination of brands and attributes (CFP). However, decision nets (Bettman, 1970) typically assume processing by brands (CPB). An alternative model, the elimination by aspects model by Tversky (1972), suggests that processing by attributes is used. The findings of this study suggest individual differences, but support the Tversky position. The discussion of acquisition effort above makes this conclusion more tentative. The task of this study would be more conducive to attribute processing than many real world brand choice situations. In actual brand choices, more processing by brands strategies may be used, because of the way supermarket shelves are organized. These real world information displays add to the difficulty of the choice task.

The distinction between brand and attribute processing may be affected by more than the task environment. Payne (1976) distinguishes between two possible ways of representing an attribute-brand combination in a list structure in memory: brand (attribute, value) or attribute (brand, value). The first representation, where the "tag" for the information is the brand, would facilitate brand processing. The second would make attribute processing easier. This issue points to a problem in coordinating the phases of problem solving in brand choice. Perceptual encoding and memory storage may be more efficient if done by brand, given typical information displays. However, as discussed below, brand comparisons may be much more effective if done by attribute. In fact there is some evidence that coding by object is more effective than encoding by attribute for accuracy of encoding (Haber, 1964; Lappin, 1967). See Kahneman (1973, especially Chapter 6) for further discussion.

Finally, the large proportion of Choice by Brand Name subjects, 20% (possibly product-class dependent), supports the findings of Jacoby, Szybillo, and Busato-Schach (1974) regarding the importance of brand name as a chunk--an organized, familiar unit of data that summarizes a great deal of information (Simon, 1974). A

<sup>4</sup>This explanation is bolstered by the fact that very few cards chosen were repeats (7.6% of all cards chosen). This implies that evaluation processes relied on retrieval from memory, while acquisition was the main process underlying the observable card sequences.

brand name itself comes to summarize configurations of attributes rather than serving as a pointer in memory for examining a detailed attribute list.

#### Implications for Designing Consumer Information Environments

In a study of unit pricing, Russo, Krieser, and Miyashita (1975) found that the typical unit price display (tags below each size of each brand on the shelf facing) makes severe information processing demands upon the consumer. They argue that a consolidated display listing unit prices for all brands and sizes is much more easily processed, and show that more consumers actually used unit price information when it was displayed in the store in such a manner. Russo (1975) extends this analysis to displays involving both price and quality information.

Theoretically, the argument underlying these findings is that comparisons of brands along a given attribute are relatively easy to make, because initially weightings across attributes (possibly incommensurate) are not required. One can determine which alternative is "best" on attribute 1 and which is "best" on attribute 2 without weighting attributes 1 and 2. Since some alternatives will be discarded just by examinations within an attribute, a smaller set of alternatives is left for making trade-off decisions. In addition, some attributes may be ignored if all brands are equivalent for that attribute. However, processing by brands involves immediate weighting (explicit or implicit) of attributes to develop an overall evaluation for each brand. Thus, processing by attributes can be argued to be cognitively easier than processing by brands in making brand comparisons for typical consumer information environments (Russo, 1975; Russo and Rosen, 1975; Tversky, 1969). In the present study, more people used attribute processing. Use of brand processing may still be relatively high because of the brand name and the information it provides. Also, Russo's arguments apply to the brand comparison phase of brand choice. For perceptual encoding and storage of information about the alternatives, brand processing may be easier, as argued above. If this conjecture is true then there may be a conflict between memory storage form (by brand) and ability to make comparisons and evaluations (easiest if organized by attribute). The consolidated display should eliminate much of the need for retrieval from memory, so that facilitation of processing by attributes would still be the most effective strategy. Memory is partially decoupled from evaluation, eliminating some of the potential conflict between storage and processing.

Thus, information should be presented in a fashion which facilitates processing by attributes. This is contrary to most store displays, which are arranged by brand. This argues for a centralized summary display to facilitate attribute processing. There is one additional factor, however. Given the evidence for the limited size of short term memory (Miller, 1956; Simon, 1974), anything which helps the consumer form chunks of information would be potentially helpful. Russo (1975) argues for more use of summary information in comprehensive point of purchase displays, and for use of small numbers (2 to 3) of such summary attributes.

Thus, in general the notion is to combine into a single display features which facilitate making brand comparisons and within attribute processing and which also

summarizes information for the consumer.<sup>5</sup> This is an instance of a more general strategy for design of information environments. This strategy is to first determine the information processing abilities and strategies of the population of interest, e.g., the form in which information is stored in memory, what kinds of information integration rules are used, the capabilities of short term memory. These parameters then are used to determine how the task of gathering information can be structured so as to be maximally congruent with the known capacities of the population. The current consumer information environment, as typified by supermarket shelf displays, seems almost maximally incongruent with what is being learned about human information processing capabilities (Russo, Krieser, and Miyashita, 1975). The typical display is set up for brand processing, the most difficult cognitive processing for comparing and evaluating brands, and summary information is typically not provided. Thus, concentration on organization of information and mode of presentation is crucial for aiding consumers through public policy (cf. Jacoby, 1974), and must augment the current focus on what specific types of information should be provided.

<sup>5</sup> Brand comparisons may not be the only element of interest. Evaluation of an entire product class relative to some standard, say for nutrition, may also be desired. However, a design which facilitates the ability of consumers to process the information presented should also suffice for this purpose. In fact, processing by attribute emphasizes the value of the entire class for each attribute.

## References

- Carol A. Berning and Jacob Jacoby, "Patterns of Information Acquisition in New Product Purchases," Journal of Consumer Research, 1(September 1974),18-22.
- James R. Bettman, "Information Processing Models of Consumer Behavior," Journal of Marketing Research, 7(August 1970).
- Bobby J. Calder, "The Cognitive Foundations of Attitudes: Some Implications for Multi-Attribute Models," in M. J. Schlinger, ed., Advances in Consumer Research, Volume II. (Chicago, Ill.: Association for Consumer Research, 1975,241-7).
- Ralph N. Haber, "Effects of Coding Strategy on Perceptual Memory," Journal of Experimental Psychology, 68(November 1964),357-62.
- G. David Hughes and Michael L. Ray, Buyer/Consumer Information Processing (Chapel Hill, North Carolina: University of North Carolina Press, 1974).
- Jacob Jacoby, "Consumer Reaction to Information Displays: Packaging and Advertising," in Salvatore F. Divita, ed., Advertising and the Public Interest (Chicago: American Marketing Association, 1974,101-18).
- Jacob Jacoby, Robert W. Chestnut, Karl C. Weigl, and William Fisher, "Prepurchase Information Acquisition: Description of a Process Methodology, Research Paradigm, and Pilot Investigation," in B. B. Anderson, ed., Advances in Consumer Research, Volume III (Chicago: Association for Consumer Research, 1976), forthcoming.
- Jacob Jacoby, George J. Szybillo, and Jacqueline Busato-Schach, "Information Acquisition Behavior in Brand Choice Situations," Purdue Papers in Consumer Psychology, No. 140, 1974.
- Daniel Kahneman, Attention and Effort (Englewood Cliffs, N.J.: Prentice-Hall, 1973).
- Joseph S. Lappin, "Attention in the Identification of Stimuli in Complex Visual Displays," Journal of Experimental Psychology, 75(November 1967),321-28.
- Denis A. Lussier and Richard W. Olshavsky, "An Information Processing Approach to Individual Brand Choice Behavior," paper presented at the ORSA/TIMS Joint National Meeting, San Juan, Puerto Rico (October 1974).
- George A. Miller, "The Magical Number Seven, Plus or Minus Two," Psychological Review, 63(March 1956),81-97.
- John W. Payne, "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," Organizational Behavior and Human Performance (1976), forthcoming.
- J. Edward Russo, "An Information Processing Analysis of Point-of-Purchase Decisions," Proceedings of the American Marketing Association Fall Conference (1975), forthcoming.
- J. Edward Russo, Gene Krieser, and Sally Miyashita, "An Effective Display of Unit Price Information," Journal of Marketing, 39(April 1975),11-19.
- J. Edward Russo and Larry D. Rosen, "An Eye Fixation Analysis of Multi-Alternative Choice," Memory and Cognition, 3(May 1975),267-76.
- Herbert A. Simon, "How Big is a Chunk?" Science, 183(February 8, 1974),482-88.
- Herbert A. Simon and Michael Barenfeld, "Information-Processing Analysis of Perceptual Processes in Problem Solving," Psychological Review, 76(September 1969), 473-83.
- Ola Svenson, "Coded Think Aloud Protocols Obtained When Making A Choice to Purchase One of Seven Hypothetically Offered Houses: Some Examples," unpublished paper, (University of Stockholm, 1974).
- Amos Tversky, "Elimination by Aspects: A Theory of Choice," Psychological Review, 79(July 1972),281-99.
- Amos Tversky, "Intransitivity of Preferences," Psychological Review, 76(January 1969),31-48.
- Peter Wright, "The Use of Phased, Noncompensatory Strategies in Decisions Between Multi-Attribute Products," Research paper 223, Graduate School of Business, (Stanford University, 1974).

## HEURISTIC SEARCH PROCESSES IN DECISION MAKING

John W. Payne, University of Chicago

### Abstract

Research in psychology indicates that individuals employ heuristics in performing a variety of complex tasks. Evidence is presented that consumers also utilize heuristics to reduce the amount of information searched in making a decision. Aspects of theory and methodology of cognitive psychology which should prove valuable to the study of consumer behavior are illustrated.

The purpose of this paper is to present evidence that consumers employ heuristic processes as a way of reducing the amount of information they have to search and evaluate in making a decision. The use of heuristic processes in the performance of complex tasks has been found in studies of such different areas of psychology as human problem solving, visual perception, and decision making under risk. In the first part of this paper, I will briefly review some of that research. The results of a recent study that examined the use of heuristic search strategies in a more consumer-oriented decision task will then be presented. Finally, I will present an example of a contingent information processing model based on the behavior revealed in that study. An underlying theme of this paper is the need for a closer integration of research in consumer behavior with the theory and methodology of more traditional areas of cognitive psychology.

### Heuristic Processing

Perhaps the most important generalization to come out of efforts to study human information processing is that an individual is a limited information processing system (Newell and Simon, 1972). In particular, the active processing of information occurs in a memory of limited capacity, duration, and ability to place information in more permanent storage. As a result, people appear to utilize heuristics that serve to keep the information processing demands of a task within the bounds of their limited cognitive capacity. Heuristic processes can be defined as

problem-solving methods which tend to produce efficient solutions to difficult problems by restricting the search through the space of possible solutions. The restriction on search is based on evaluation of the structure of the problem (Braunstein, 1976).

The same restriction on search which increases efficiency may, at times, result in individuals ignoring or misusing information in reaching a judgment or achieving a solution to a problem. Heuristic processes, in other words, are procedures used by individuals which sacrifice the certainty of a correct judgment for increased efficiency in the process.

Empirical support for the concept of heuristic processing has been found in studies in a wide variety of areas of cognitive psychology. For example, clear illustrations of the use of heuristic processes by individuals can be found in studies of human problem solving (Newell & Simon, 1972), pattern recognition, depth perception (Braunstein, 1976), probability estimation, and

risky decision making. In the next two sections, I will briefly review some of the evidence for heuristic processing in two of these areas of cognitive psychology. More intensive discussions of the use of heuristics by individuals may be found in Newell and Simon (1972) and Braunstein (1976).

### Probabilistic Information Processing

Some of the most striking evidence for the use of heuristics by individuals has been provided by Kahneman and Tversky in their studies of probabilistic information processing. In the case of probability estimation, several recent studies have demonstrated that humans seem to follow a judgmental heuristic called "representativeness" when faced with tasks involving intuitive prediction (1972, 1973). The idea is that people predict the outcome that appears most representative of the evidence. While this heuristic often leads to correct judgments, it can lead to large and consistent biases that are quite difficult to eliminate. The reason is that there are factors such as the reliability of evidence and the prior probability of the outcome which affect the likelihood of the outcome but not its representativeness. For example, Kahneman and Tversky (1973) have given subjects a brief personality description allegedly sampled at random from a group of 100 professionals--engineers and lawyers. In contrast to the normative expectation, the subjects rated the probability of that person being an engineer rather than a lawyer the same regardless of whether they were told that the group consisted of 70 engineers and 30 lawyers, or 30 engineers and 70 lawyers. Apparently, subjects considered the degree to which the descriptions were representative of the two stereotypes, lawyers or engineers, with relatively little regard for prior probabilities in making their judgments.

Tversky and Kahneman (1974) have discussed two other heuristics which appear to be important in probabilistic information processing. The first is a heuristic called "availability." According to the availability heuristic, one judges "the frequency of a class or the probability of an event by the ease with which instances or occurrences can be brought to mind" (p. 1127). Slovic, Fischhoff, and Lichtenstein (1976) report an experiment illustrating how the availability heuristic may lead to important biases in people's perceptions of low-probability, high-consequence events. Bias may also occur when people attempt to ease the strain of processing information through the use of a heuristic called "anchoring and adjustment." This heuristic operates by an individual selecting a natural starting point or anchor in the task environment to be used as a first approximation to the judgment. This anchor is then adjusted to accommodate the implications of additional information. Work by Tversky and Kahneman (1974) and Lichtenstein and Slovic (1971) demonstrates that typically the adjustment is crude and imprecise and fails to do justice to the importance of additional information.

### Decision Making Under Risk

Research on how people choose among alternative gambles provides additional evidence that individuals tend to use heuristic processes. The use of heuristics in risky decision making is especially interesting since



There exist rather straightforward computational algorithms which guarantee an optimal solution. Risky decision behavior is also highly relevant to consumer choice behavior. Both involve decisions about multi-dimensional alternatives. In addition, there is a substantial literature relating perceived risk to consumer behavior (Ross, 1975).

In the typical risky decision making task, an individual is faced with two alternative gambles. The probability of winning, probability of losing, amount to be won, and amount to be lost are displayed for each gamble. In such situations, people appear to choose among gambles on the basis of a direct comparison among these displayed values, rather than on the basis of maximizing expected value. While choices based on the heuristic use of the displayed probabilities and amounts usually correlate well with the choices which would be obtained with expected value maximization, or at least with the maximization of a function of subjective transformations of probability and value, there are situations in which these heuristics lead to preferences which fail to correspond to those expected from either objective or subjective maximization functions. It has been shown, for example, that people will exhibit consistent preferences among gambles which differ in displayed probabilities and amounts, even when the gambles are mathematically identical and the choice of one over the other can have no objective consequence (Payne and Braurstein, 1971). It has also been shown that the use of displayed probabilities and amounts is not merely due to ignorance of alternative procedures or difficulty in making the required computations. These tendencies persist even after subjects are instructed in the meaning of expected value and given the expected value of each alternative gamble (Lichtenstein, Slovic, and Zink, 1969). Additional evidence for the use of heuristics in risky decision making may be found in studies by Lichtenstein and Slovic (1971), Payne (1975), and Tversky (1969, 1972).

An example of a model of risky decision behavior based on heuristic processes was proposed by Payne and Braurstein (1971). This model is of interest, in part, because it assumes that individuals may search and utilize only part of the information available about the alternative gambles. The hypothesized decision strategy was presented in the form of a process model involving two stages: an evaluation stage followed by a choice stage. The particular choice rule used by the decision maker was assumed to be contingent upon the outcome of the evaluation process. Specifically, the model assumed that the decision maker initially considers the probability relationship within the gambles in a pair. If the probability to win is less than the probability to lose, a choice rule based on selecting the gamble with the lesser probability to lose is evoked. If the probability to win is greater than the probability to lose within each gamble in the pair, then an attempt is made to maximize the amount to win. In the case where the amounts are equal within a pair, the probability to win may be used as a secondary criterion, and the gamble with the greater probability to win chosen. For pairs of gambles where the probability to win equals the probability to lose, there are alternative paths which can be taken in the model, possibly corresponding to two types of subjects. These types of pairs of gambles may be treated as if they were pairs of gambles in which the probability to win is greater than the probability to lose, or as if they were pairs of gambles in which the probability to win was less than the probability to lose. While the results that this model was developed to explain were obtained in a study using special types of pairs of gambles, the form of the model provides a clear illustration of a heuristic decision strategy. Based upon an evaluation of the structure of the problem, the

decision maker is assumed to restrict his search and use of information to just one or two dimensions of the gambles in making his decision.

The research reviewed here in probabilistic information processing and risky decision making is consistent with the generalization reached by Slovic and Lichtenstein (1971) in their comprehensive review of Bayesian and regression studies of human judgment:

We find that judges have a very difficult time weighting and combining information--be it probabilistic or deterministic in nature. To reduce cognitive strain, they resort to simplified decision strategies, many of which lead them to ignore or misuse relevant information (p. 115).

In terms of problem solving, Simon and Newell (1971) have proposed that the central process in human problem solving is the use of heuristic methods to carry out highly selective searches of problem spaces.

### Consumer Decision Behavior

The processing of information by consumers has recently received increased attention. One finding of this research is that "the amount of information sought is typically small relative to the amount of information available" (cf. Jacoby, 1975). Haines (1974) has advanced a "Principle of Information Parsimony." The idea is that "consumers seek to process as little data as is necessary in order to make rational decisions" (p. 96). This principle seems entirely consistent with results obtained in other areas of cognitive psychology. Although the work of Simon (1957) suggests that the term "rational" be replaced by the term "satisfactory," the interesting question then becomes "how do decision makers go about the process of reducing the amount of information sought in order to make a decision?" Haines (1974) suggests a partial answer when he argues that "it is important to understand that people do take advantage of patterns in the task environment to reduce information processing" (p. 96). The definition of a heuristic procedure in problem solving given earlier incorporates the similar idea that the restriction of search through the space of possible solutions will be based on an evaluation of the structure of the problem (Braurstein, 1976). It is the central thesis of this paper that consumers utilize heuristic strategies in making decisions as a way of reducing the amount of information they must search and evaluate.

In the next section of the paper, I will briefly review the results of a recent study which show that individuals, when faced with the problem of choosing a preferred alternative in a sufficiently complex situation, will adopt heuristic decision strategies which produce highly selective searches among the sets of possible alternatives. An example of a computer-based process model that attempts to simulate the behavior revealed in the study will then be presented. The model illustrates the use of a heuristic decision procedure to simplify a decision task.

### Contingent Decision Making

Research on problem solving has shown individual behavior to be highly adaptive to the demands of the task. This suggested that the information processing procedures used by decision makers might be systematically related to certain characteristics of the decision situation. The study discussed below used two process tracing techniques: (1) explicit information search and (2) verbal protocols, to make clear the effects of

two variations in the complexity of a decision task; (1) number of alternatives and (2) number of dimensions, on the information processing strategies subjects use in reaching a preferential choice. It was hypothesized that increases in the complexity of a decision situation would result in decision makers resorting to choice heuristics in an effort to reduce cognitive strain. Two experiments were conducted. (The interested reader is referred to Payne, 1976 for a complete report on this study.)

#### Method

The method in both experiments was very similar. The second experiment was an extension and replication of the first experiment. Six subjects who were paid college students participated in the first experiment (12 subjects in the second). The stimuli were "information boards" representing different one bedroom furnished apartments. An information board consisted of a number of envelopes containing cards labeled with the name of a dimension (attribute), e.g., "noise level." To obtain the value of that dimension for a particular alternative, the decision maker had to pull the card out of the envelope, turn it around and place it back into the envelope. The information about the value of the dimension was on the back of the card, e.g., "noise level--low." Once a card was turned over, the value of the dimension on the particular alternative was clearly displayed for the remainder of the choice problem.

A decision situation for a subject involved a number of alternatives, either 2, 6, or 12 (2, 4, 8, or 12 in the second experiment), and a number of dimensions of information available per alternative, either 4, 8, or 12. There were three levels of value on such dimension. The dimensional values of an alternative were selected so that each alternative would have a priori both good and poor qualities.

The procedure was very simple. Each subject was run individually in one hour sessions. The subjects were told that they would be presented with a number of alternatives to choose among and a certain amount of information about each alternative. The subject was not instructed in how much of the information he or she had to use in reaching a decision. No time constraints were placed on the subjects. They were instructed to work at their own pace and that they should have plenty of time to finish.

In addition to instructions regarding the use of the information boards, the subjects were also instructed to "think aloud" while making their decision. Verbal protocols have proved useful as the basis for models of human behavior in the performance of a variety of tasks (cf. Newell & Simon, 1972).

#### Results (First Experiment)

A complete transcript of the verbal reports given by each subject was made. Consistent with the procedure suggested by Newell and Simon (1972), the protocols were broken up into short phrases. During the analysis of the results, excerpts from the protocols of the six subjects will be presented. The complete protocols may be obtained from the author.

The search data for each subject was primarily organized in terms of amount of available information search per alternative in a choice set and the pattern of search, interdimensional or intradimensional. The pattern of search was determined by examining the alternative and dimension associated with the  $n$ th + 1 piece of information searched by a subject as a function of the  $n$ th piece of information search. Details on how the pattern of search was determined are available in

Payne (1976). Similar analyses of single-step transition search data can be found in Bettman and Jacoby (1975).

These two characteristics of the search data, amount of information searched per alternative and pattern of search, proved useful in discriminating between four alternative models of decision making which were of particular interest in this study. The four models were: additive, additive difference, conjunctive, and elimination-by-aspects. Each of these models implies, at least in its most common forms, different information search processes.

An additive decision strategy implies an interdimensional pattern of search and a constant amount of search per alternative. The constant pattern of search is implied because the additive model assumes that each alternative in a choice set is evaluated separately and that the comparison process involves only the overall values that have been determined for each alternative. A decision maker following a strict additive difference model would also have to search a constant amount of information per alternative, but would search in an intradimensional fashion. Both the conjunctive and elimination-by-aspects models, on the other hand, imply the possibility of a decision maker using a variable search pattern. For example, the elimination-by-aspects model (Tversky, 1972), describes choice as a covert elimination process. In choosing among multidimensional alternatives, the individual is assumed to proceed in the following manner. A dimension or aspect is selected. Then all the alternatives that do not possess that dimension or aspect are eliminated. Consequently, those alternatives eliminated early would have only a few (one) dimension searched. The elimination procedure is repeated until all but one of the alternatives are eliminated. Those alternatives eliminated late in the choice process would have more dimensions searched. A conjunctive decision strategy would also result in a pattern of search involving a minimum of one dimension examined on some alternatives (those unsatisfactory on the initial dimension searched) up to an examination of all the relevant dimensions of the preferred alternative. Both the conjunctive and elimination-by-aspects processes are examples of heuristics by which people seek to reduce the amount of information processing involved in complex decision making. The two models differ, however, in whether they imply an interdimensional (conjunctive) or intradimensional (elimination-by-aspects) search strategy.

The information search pattern exhibited by each subject was determined for each of the different decision situations. The results indicated that the information processing leading to a preferential choice will vary as a function of task complexity. The most important determinant of complexity examined was clearly the number of alternatives available. Table 1 shows the classification of the search patterns for each subject as a function of the number of alternatives available.

It is clear from Table 1 that when faced with a more complex decision task, either six or twelve alternatives, subjects employed decision strategies which resulted in a variable amount of information searched across alternatives. This supports the hypothesis that increases in the complexity of a decision situation will result in decision makers resorting to choice heuristics in an effort to reduce cognitive strain. As mentioned, the use of an elimination-by-aspects choice process by a decision maker represents an example of such a heuristic.

The verbal protocols obtained provide further evidence that subjects tended to adopt decision strategies which would eliminate some of the available alternatives as

TABLE 1  
CLASSIFICATION OF SEARCH PATTERNS AS A FUNCTION  
OF NUMBER OF ALTERNATIVES AVAILABLE

Number of Alternatives	Search Pattern			
	Interdimensional		Inradimensional	
	Constant	Variable	Constant	Variable
2	Subject 1 Subject 6		Subject 2 Subject 3 Subject 4 Subject 5 <sup>a</sup>	
6		Subject 1 Subject 5		Subject 2 Subject 3 Subject 4 <sup>b</sup> Subject 6 <sup>b</sup>
12		Subject 1 Subject 4 Subject 5		Subject 2 <sup>b</sup> Subject 3 <sup>b</sup> Subject 6 <sup>b</sup>

<sup>a</sup>This subject only examined one piece of information, i.e., rent, in choosing among the two alternatives.

<sup>b</sup>A mixed search strategy which initially involved intradimensional search and ended with interdimensional search.

quickly as possible and on the basis of a limited amount of information search and evaluation. For example, consider this excerpt from the protocol of subject 1:

- A162: Apartment E.  
A163: The rent for apartment E is \$140,  
A164: which is a good note.  
A165: The noise level for this apartment is high.  
A166: That would almost deter me right there.  
A167: Ah, I don't like a lot of noise.  
A168: And, if it's high, it must be pretty bad.  
A169: Which means, you couldn't sleep.  
A170: I would just put this aside right there. I wouldn't look any further than that.

Another explicit example of the elimination of alternatives after a limited search and evaluation is provided by this example from the protocol of subject 4:

- D289: Since we have a whole bunch here,  
D290: I'm going to go across the top and  
D291: see which noise levels are high.  
D292: If there are any high ones,  
D293: I'll reject them immediately.  
:  
D297: Go to D.  
D298: It has a high noise level.  
D299: So, we'll automatically eliminate D.

#### Results (Second Experiment)

The results obtained in the second experiment were similar to those obtained in the first experiment. The conclusion that as the number of available alternatives increases, decision makers shift from decision strategies involving a constant amount of search per alternative, e.g., compensatory procedures, to decision strategies which involved eliminating some alternatives on the basis of only a few dimensions, e.g., conjunctive or elimination-by-aspects procedures, was supported. A three-way analysis of variance (number of

alternatives available, amount of information available per alternative, and subjects) showed that the amount of variation in percentage of available information searched increased as the number of alternatives increased,  $F(3,33) = 12.36$ , ( $p < .01$ ). The main effect of amount of information per alternative,  $F(2,22) = .78$ , and the interaction,  $F(6,66) = 1.10$ , were not significant.

Stronger support for the use of a strict elimination-by-aspect process by decision makers was obtained by calculating the number of search patterns that showed not only a variable and intradimensional pattern but also a pattern where all (remaining) alternatives would be completely searched on a dimension before some subset of those alternatives would be searched on another dimension. Thirty-three of the 108 multi-alternative search patterns were consistent with a strict elimination-by-aspects decision process. For 99 of the 108 multialternative choice situations (nine for each subject), the amount of available information searched was as great as or greater for the alternative chosen than for any other alternative in the choice set. In the remaining nine situations, the preferred alternative had next to the maximum amount of information searched.

#### Discussion

It is clear from the results of the two experiments just discussed that individuals will employ decision strategies resulting in a restricted pattern of information search when the choice task becomes sufficiently complex. These decision strategies can be characterized as heuristic processes similar to those found in studies of human problem solving. Evidence for the use of heuristics has been found in another recent study of information seeking behavior in multidimensional choice (Russo & Doshier, 1975).

Important individual differences in information processing also were shown to exist in both the experiments. In particular, some of the subjects tended to search for information in an interdimensional fashion, others in an intradimensional fashion. Bettman and Jacoby (1975) also found such individual differences.

It was suggested in Payne (1976) that one possible explanation for differences in information search may be in how the decision maker represents the knowledge he acquires about the alternatives in the decision task. For example, a decision maker might store information about the decision alternatives in the following form: Apartment A (rent, \$140). This object (property, value) representation would suggest that an individual might find it easier to search and evaluate (store) information within an alternative and across dimensions. On the other hand, a decision maker might choose to code information in terms of rent (apartment A, \$140). This property (object, value) representation would suggest that an intradimensional form of processing might be easier for an individual. This type of knowledge representation has been widely used in building theories of human cognitive processes (Simon and Newell, 1974). In the context of consumer choice, Calder (1975) has also pointed out that the structure of decision processes may depend upon how information is coded in memory. Bettman and Jacoby (1975) present a further discussion of the implications of individual differences in memory storage for consumer information processing. Research aimed at relating an individual's pattern of search to any preferences that individual might have for encoding information in memory is currently being undertaken.

The results of the information search experiments presented above indicate: (1) that the same individual will process information in different ways as a function of simple task variations; and (2) that different individuals will process information in different ways in the same decision task. Both these findings support conclusions reached by Haines (1974). However, there was also evidence of different processes being used by the same individual within a single decision situation. Table 2a presents excerpts from the protocol of subject 2 in the first experiment when faced with a 12 alternative choice situation.

TABLE 2  
PROTOCOL AND MODEL TRACE FOR SUBJECT 2

(A) Protocol	(B) Trace
Let's just see what the rents are in all the apartments first.	I WILL NOW LOOK AT THE RENT OF EACH ALTERNATIVE.
The rent of A is \$140. The rent of B is \$110. The rent of C is \$170. The rent of D is \$110. The rent of E is \$170. F is \$170. G is \$170. H is \$110. I is \$110. J is \$140. K is \$140. L is \$140.	THE RENT OF A IS \$140. THE RENT OF B IS \$110. THE RENT OF C IS \$170. THE RENT OF D IS \$110. THE RENT OF E IS \$170. THE RENT OF F IS \$170. THE RENT OF G IS \$170. THE RENT OF H IS \$110. THE RENT OF I IS \$110. THE RENT OF J IS \$140. THE RENT OF K IS \$140. THE RENT OF L IS \$140.
Um, \$170 is too much. But, if the other ones aren't good, I'll look at them later. But right now I'll look at the other ones. ...	ALTERNATIVE C ELIMINATED. ALTERNATIVE E ELIMINATED.  ALTERNATIVE F ELIMINATED. ALTERNATIVE G ELIMINATED.  ...
I'm going to look at landlord attitude.  In H it's fair. In D it's poor. B it's fair, and A it's good. In L the attitude is poor. In K it's poor. In J it's good, and in I it's poor.  So, one of them . . . is poor. So that's important to me. ...	I WILL NOW LOOK AT THE LANDLORD OF EACH ALTERNATIVE.  THE LANDLORD OF A IS GOOD. THE LANDLORD OF B IS FAIR. THE LANDLORD OF D IS POOR. THE LANDLORD OF H IS FAIR. THE LANDLORD OF I IS POOR. THE LANDLORD OF J IS GOOD. THE LANDLORD OF K IS POOR. THE LANDLORD OF L IS POOR.  ALTERNATIVE D ELIMINATED.  ALTERNATIVE I ELIMINATED. ALTERNATIVE K ELIMINATED. ALTERNATIVE L ELIMINATED.  ...
Kitchen facilities in A are poor. In B poor. In J fair, and in H they're good.  Oh, J and H have better kitchen facilities than A and B.	I WILL NOW LOOK AT THE KITCHEN OF EACH ALTERNATIVE. THE KITCHEN OF A IS POOR. THE KITCHEN OF B IS POOR. THE KITCHEN OF H IS GOOD. THE KITCHEN OF J IS FAIR.  ALTERNATIVE A ELIMINATED. ALTERNATIVE B ELIMINATED.

(A) Protocol	(B) Trace
And everything else is about the same. So eliminate those two. And, decide between these two.  ...	SINCE I'VE ONLY GOT TWO ALTERNATIVES I WILL APPLY A COMPENSATORY DECISION PROCESS. THE CHOICE IS BETWEEN H AND J.  ...
Let's see furniture quality. In H is below average. In J it's below average so that's about the same there.  ...	LET'S SEE ON FURNITURE. H IS BELOW AVERAGE. J IS BELOW AVERAGE. H AND J ARE EQUAL ON THAT.  ...
Landlord attitude in J is better than in H. ...	LET'S SEE ON LANDLORD J IS BETTER  LET'S SEE ON SIZE J IS BETTER
The rents are just about the same. In both of them cleanliness is poor. In J the rooms are larger, so, I guess, J will be better.	LET'S SEE ON RENT H IS BETTER  I CHOOSE ALTERNATIVE J.

The protocol clearly shows this decision maker initially using a strict elimination-by-aspects process to eliminate alternatives. Support for this conclusion was also obtained from a detailed analysis of the single-step transition data in the individual's pattern of search. Notice how this decision maker appears to reduce the choice problem from 12 alternatives to eight alternatives, and eventually to just a pair of alternatives. At that point, the protocol shows the decision maker shifting from an elimination-by-aspects procedure to what appears to be an additive difference strategy. Eventually, after directly comparing the final two choice alternatives, the decision maker chose apartment J. Other subjects also give indications of similar combinations of decision processes. Einhorn (1971) suggested a similar explanation of decision making in complex situations. See also Wright (1974).

One way to conceptualize the four decision processes examined in the previous two studies would be as different subroutines in a general choice program. The control conditions under which one of these sets of processes might be called would seem to depend, at least in part, on the characteristics of the decision problem. In that respect, the less cognitively demanding decision procedures, conjunctive and elimination-by-aspects, might be called early in the decision process as a way of simplifying the decision task by quickly eliminating alternatives until only a few alternatives remained as choice possibilities. The decision maker might then employ one of the cognitively more demanding choice procedures, e.g., additive difference model, to make the final evaluations and choice.

Model

A contingent process model that attempts to simulate the behavior revealed by subject 2 was developed. The model is based on assumptions derived from the analysis of the verbal report of subject 2 (excerpts of that

report are given in Table 2a). The model encompasses three categories of behavior: (1) a control process that selects one of two decision procedures on the basis of an evaluation of the complexity of the decision task, (2) an elimination of alternatives procedure, and (3) a compensatory comparison procedure:

dimension (D). Else respond alternatives equal and go to D2.

D5: Compare overall worth of alternative (X) with overall worth of alternative (Y).

If worth of alternative (X) greater, then respond alternative (X) preferred. Else if worth of alternative (Y) greater, then respond alternative (Y) preferred. Else R $\emptyset$ .

R $\emptyset$ : Respond--No choice possible.

The model has been coded in BASIC. The output (trace) of the program corresponding to the parts of the verbal protocol given in Table 2a is presented in Table 2b.

There are a number of ways of comparing the behavior of an individual with the output of a computer-based model that attempts to simulate that behavior. At the summary level of final choice, the model presented here selected the same alternative out of the 12 possible that the decision maker selected. This is, of course, the first and most necessary property of such a model. However, the real value of a process (computer) model is in the ability of the model to account for the sequential aspects of the decision maker's behavior.

An analysis of the verbal protocol of subject 2 indicated four major types of statements: (1) goal statements, e.g., "Let's just see what the rents are in all the apartments first," (2) statements reflecting search, e.g., "The rent of A is \$140," (3) evaluation statements, e.g., "Um, \$170 is too much," and (4) comparison statements, e.g., "In J the rooms are larger." The trace shown in Table 2b also contains four types of statements which correspond to those contained in the subject's protocol. A comparison of the sequence of goal statements, observations, evaluative statements, and comparative statements contained in the protocol and trace presented in Table 2 suggests that the correspondence between the subject's behavior and the model is reasonable. More formally, it is possible to construct Problem (Decision) Behavior Graphs (Newell and Simon, 1972) based on the decision maker's protocol and on the trace of the model. One way to analyze the correspondence between the two Decision Behavior Graphs would be in terms of the sequential dependencies among the various types of statements (behavior). That is, how often was one type of behavior followed by another type of behavior. This analysis is based on a suggestion for the analysis of production systems presented by Newell (1966). Table 3 presents the results of such a sequential analysis. The model appears capable of providing sequential behavior which corresponds closely to the actual sequential behavior of the subject.

Unfortunately, space limitations make it impossible to expand upon this comparison of the model trace with the subject's protocol. A more intensive treatment of the protocol and model presented here would, for example, give a detailed examination of the exact information being searched and evaluated at each point in time in both the protocol and the trace. However, it should be noted that the heuristic processing form of this model is consistent with research in other areas of behavior (cf. Newell and Simon, 1972). Examples of similar types of process models dealing with consumer choice are provided by Bettman (1970) and Haines (1974). For a more complete discussion of the role of verbal protocol analysis in decision making research, see Payne (1974). See also Bettman (1974) for several interesting suggestions regarding procedures for formally analyzing simulation models.

- C $\emptyset$ : Evaluate complexity of decision task by determining number of alternatives in List-available alternatives.
- C1: If the number of alternatives is greater than two then E $\emptyset$ .
- C2: If the number of alternatives is exactly two then D $\emptyset$ .
- C3: If the number of alternatives is less than two then R $\emptyset$ .
- E $\emptyset$ : Eliminate alternatives by dimensions.
- E1: Search List-dimensions for first (next) most important dimension (D), and set goal--Eliminate by dimension (D), if end of list then R $\emptyset$ .
- E2: Search List-available-alternatives for first (next) alternative (A), if end of list then C $\emptyset$ .
- E3: Search external environment for dimension (D, value) for alternative (A).
- E4: Search memory for List-acceptable-dimension (D, values).
- E5: Determine if dimension (D, value) for alternative (A) on List-acceptable (D, values). If value on List then E2, else mark alternative (A) eliminated and remove from List-available-alternatives.
- E6: Go to E2.
- D $\emptyset$ : Direct Comparison Process.
- D1: Search List-available-alternatives for first alternative (X) and next alternative (Y), and set goal--comparison (X) and (Y).
- D2: Search List-dimensions for last (next) most important dimension--D, if end of list then D5.
- D3: Search memory for dimension (D, value) for alternative (X) and alternative (Y), if not available then search external environment for dimension (D, value) for alternative (X) and for alternative (Y).
- D4: Compare dimension (D, value) for alternative (X) and dimension (D, value) for alternative (Y) with ordered List-acceptable dimension (D, values).
- If alternative (X) ordered higher than alternative (Y) on dimension (D), then respond alternative (X)--better, and increment overall worth of alternative (X) by relative importance value of dimension (D). Else if alternative (Y) ordered higher than alternative (X) on dimension (D), then respond alternative (Y) better, and increment overall worth of alternative (Y) by relative importance value of

TABLE 3  
THE NUMBER OF TIMES THAT EACH TYPE OF STATEMENT (ROWS) WAS FOLLOWED BY EACH TYPE OF STATEMENT (COLUMNS)

Type of Statement	Type of Statement			
	Goal	Search	Evaluation	Comparison
Goal	2 <sup>a</sup> (1)	10 (10)	∅ (∅)	6 (8)
Search	5 (5)	43 (42)	3 (3)	2 (2)
Evaluation	3 (3)	∅ (∅)	2 (∅)	1 (∅)
Comparison	7 (9)	∅ (∅)	1 (∅)	1 (∅)

<sup>a</sup>This number is taken from the subject's protocol. The number below in parentheses is taken from the trace of the model.

#### Conclusion

The emphasis in this paper has been on relating the study of consumer information search and decision making to certain theoretical concepts and methodological techniques from areas of cognitive psychology, such as human problem solving. Of particular concern was applying the concept of heuristic processes to explain the limited amount of information search and evaluation consumer decision makers appear to employ when faced with complex problems. Evidence supporting the use of heuristic decision strategies was found in a review of studies of complex judgment in two areas of cognitive psychology closely related to consumer behavior, and in the results of two experiments specifically aimed at investigating information search in a complex decision task. It appears that consumers do utilize heuristics to reduce the amount of information they have to process in making a decision.

#### References

James R. Bettman, "Information Processing Models of Consumer Behavior," Journal of Marketing Research, 7 (August, 1970), 370-376.

. "Toward a Statistics for Consumer Decision Net Models," Journal of Consumer Research, 1 (June, 1974), 71-80.

James R. Bettman and Jacob Jacoby, "Patterns of Processing in Consumer Information Acquisition," Working Paper No. 31, Center for Marketing Studies, University of California, Los Angeles, 1975.

Myron L. Braunstein, Depth Perception Through Motion (New York: Academic Press, 1976).

Bobby J. Calder, "The Cognitive Foundations of Attitudes: Some Implications for Multi-Attribute Models," in M. J. Schlinger, ed., Advances in Consumer Research, Volume II (Chicago: Association for Consumer Research, 1975), 241-247.

George H. Haines, Jr., "Process Models of Consumer Decision Making," in G. D. Hughes and M. L. Ray, eds., Buyer/Consumer Information Processing (Chapel Hill, NC: University of North Carolina Press, 1974).

Jacob Jacoby, "Consumer Psychology: An Octennium," Annual Review of Psychology, 27 (1976), in press.

Daniel Kahneman and Amos Tversky, "On the Psychology of Prediction," Psychological Review, 80 (1973), 237-251.

Daniel Kahneman and Amos Tversky, "Subjective Probability: A Judgment of Representativeness," Cognitive Psychology, 3 (1972), 430-454.

Sarah Lichtenstein and Paul Slovic, "Reversals of Preference Between Bids and Choices in Gambling Decisions," Journal of Experimental Psychology, 89 (1971), 46-55.

Sarah Lichtenstein, Paul Slovic, and Donald Zink, "Effect of Instruction in Expected Value on Optimality of Gambling Decisions," Journal of Experimental Psychology, 79 (1969), 236-240.

Allen Newell, "On the Analysis of Human Problem Solving Protocols," Department of Computer Science Paper, Carnegie-Mellon University, Pittsburgh, Pa., 1966.

Allen Newell and Herbert A. Simon, Human Problem Solving (Englewood Cliffs, N. J.: Prentice-Hall, 1972).

John W. Payne, "A Process Tracing Study of Risky Decision Making: Examples of Protocols and Comments," Complex Information Processing Working Paper No. 274, Carnegie-Mellon University, Pittsburgh, Pa., 1974.

. "Relation of Perceived Risk to Preferences Among Gambles," Journal of Experimental Psychology: Human Perception and Performance, 1 (1975), 86-94.

. "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," Organizational Behavior and Human Performance (1976), in press.

John W. Payne and Myron L. Braunstein, "Preferences Among Gambles with Equal Underlying Distributions," Journal of Experimental Psychology 87 (1971), 13-18.

Ivan Ross, "Perceived Risk and Consumer Behavior: A Critical Review," in M. J. Schlinger, ed., Advances in Consumer Research, Volume II (Chicago: Association for Consumer Research, 1975), 1-19.

J. Edward Russo and Barbara Anne Doshier, "Dimensional Evaluation: A Heuristic for Binary Choice," unpublished manuscript, Department of Psychology, University of California, San Diego (August, 1975).

Herbert A. Simon, Models of Man (N.Y.: Wiley, 1957).

Herbert A. Simon and Allen Newell, "Human Problem Solving: The State of the Theory in 1970," American Psychologist, 26 (1971), 145-159.

. "Thinking Processes," in D. H. Krantz, R. C. Atkinson, R. D. Luce, and P. Suppes, eds., Contemporary Developments in Mathematical Psychology, Volume 1 (San Francisco: Freeman, 1974).

Paul Slovic, Baruch Fischhoff, and Sarah Lichtenstein, "Cognitive Processes and Societal Risk Taking," in J. S. Carroll and J. W. Payne, eds., Cognition and Social Behavior (Potomac, Md.: Lawrence Erlbaum Associates, 1976).

Paul Slovic and Sarah Lichtenstein "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment," Organizational Behavior and Human Performance, 6 (1971), 649-744.

Amos Tversky, "Elimination by Aspects: A Theory of Choice," Psychological Review, 79 (1972), 281-299.

. "Intransitivity of Preferences," Psychological Review, 76 (1969), 31-48.

Amos Tversky and Daniel Kahneman, "Judgment Under Uncertainty: Heuristics and Biases," Science, 185 (1974), 1124-1131.

Peter Wright, "The Use of Phased, Noncompensatory Strategies in Decisions between Multi-Attribute Products," Research Paper 223, Graduate School of Business, Stanford University, 1974.

This research was supported by Research Grant MH25788 from the National Institute of Mental Health.

## CONSUMER PERCEPTION OF ADVERTISING EFFECTIVENESS

### ~~—AND LIFE STYLE CONGRUENCY~~

D. W. Greeno, University of Toronto  
G. H. Haines, Jr., University of Toronto  
M. S. Sommers, University of Toronto

#### Abstract

A sample of Canadian housewives was exposed to three levels of life style advertising message congruency and then asked to rate the advertisements along a number of traditional additudinal pre-test rating scales of effectiveness. The results indicate that subjects do not rate the most congruent communication message as the most effective. In fact, the one life style "type" investigated here rated the most "incongruent" advertisement as being the most effective. On the basis of pre-test rating scales one might conclude that message-receiver congruence may not be a valid marketing communications strategy in all cases.

#### Introduction

The research tradition in consumer behavior has evolved from one that was molecular in nature to a current one that is more molar. Research focussed narrowly on such individual constructs as personality, social class, reference groups, or self image in the molecular tradition. These constructs, however, were not integrated in an operational sense until the life-style analytic format (Wells and Tigert, 1971; Tigert, 1972), a more molar approach, was developed.

However, whether the research tradition was molecular or molar, managerial prescriptions were offered. These prescriptions, while apparently reasonable, were rarely empirically validated. For example, one implication for marketing communications that has often been stated is that the message sender should reinforce, in a congruent fashion, a receiver's personality, social class, reference group, self image, or life-style profile. The marketing communications literature (Kernan, Dommermuth and Sommers, 1970) however does not directly demonstrate that message-receiver congruence is a valid communication strategy. One previous study of the congruence principle which worked with personality variables (need for achievement and affiliation) found very mixed results indeed (Whittaker, 1972). This paper reports on an experiment designed to test the validity of the message congruence prescription. The research was designed to address the question: would information be processed in a more favorable manner if the congruency rule with respect to life style is followed? That is, is the prescription of product-buyer congruency valid?

#### Theory

Even a cursory view of the task of extracting information from an advertisement would suggest that a major component of such a task is the processing of sensory, particularly visual, information. The advertisement perception and processing task can be characterized as consisting of three elements:

1. detection - the environmental information necessary to notice the advertisement must be available to the consumer. Available evidence (Ackoff and Emshoff 1975,

p. 98) suggests this is not always the case

2. diagnosis - the consumer must formulate a set of alternative actions available, and

3. decision - the consumer must select and accurately execute the optimum (in terms of the consumers' own goals) alternative within a specific time span.

This section describes briefly a theory of perceptual information processing, due to Broadbent (1971) and Sokolov (1969), which has been successfully applied to predicting effective driver decision making by Mihal (1974). This theory can be used as a conceptual basis for the identification of a number of predictions of consumer reactions to advertising, and also as a basis for explaining how these predictors function. It should be noted other theories are available, such as that proposed by Wright (1974). The justification for using the Broadbent-Sokolov theory is a previous research base (see Sokolov 1969 and Mihal 1974 and references therein) which supports its validity as a general model of perceptual information processing.

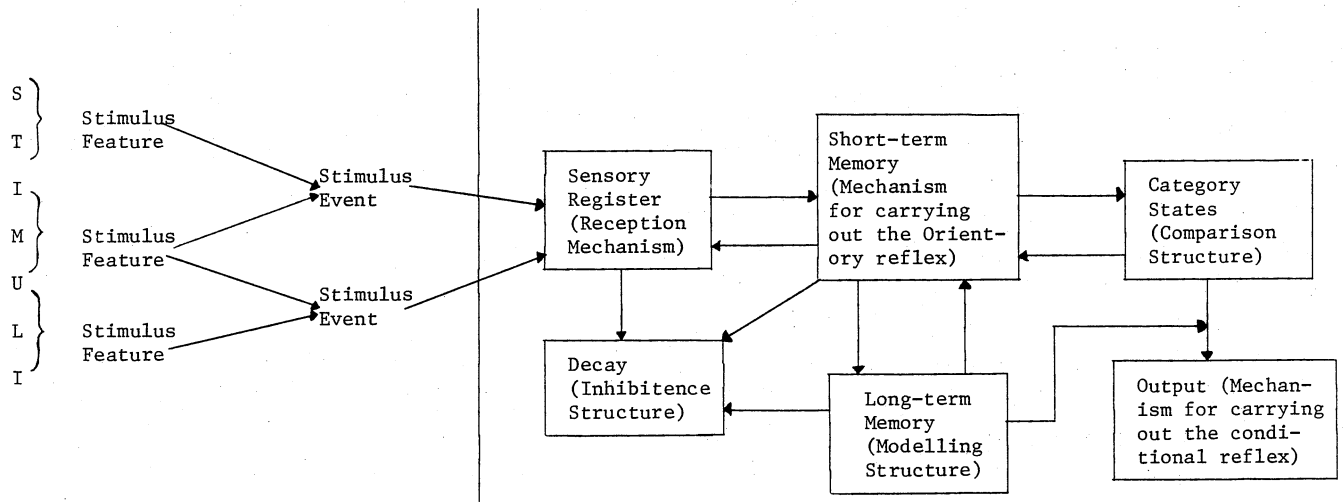
Overview of the Model. Any given possible state of the environment can be characterized by various stimulus features. Stimulus features combine systematically to produce a stimulus event. Incoming data from stimulus events are transmitted to an individual's sensory register, where the process of information analysis begins. The analysis at this point appears to be primarily on physical and/or descriptive dimensions, such as color, spatial relations, and time. Sensory data are then transferred to short-term memory. The sensory data is then further analyzed and categorized in short-term memory. Next it is held in category states which, roughly, correspond to response alternatives. Then it is transferred to long-term memory or lost from the system. Loss is strongly influenced by the limited capacity of short-term memory. Attention, stress, fatigue, visual interference, prior knowledge, and the number of perceptual categories available all influence the transfer of information into short-term memory.

Three major processes described by Broadbent as occurring in this system appear to be important. First, the system appears to continuously monitor incoming data for novelty or change. The intensity of the monitoring depends on the amount of data coming in, and on the expectation that a change will occur. Thus, if a person watching television does not expect, on the basis of past experience, to see an advertisement, it is quite possible he or she "won't even see" the advertisement even though he or she is "looking right at it". This process appears to be based primarily on an individual's behavior perceptual skills, that is, his or her ability to detect relevant data in the environment even when the event rate (i.e., the rate of incoming data) is excessive. Therefore, one crucial element at an individual level is perceptual ability, or effectiveness in separating relevant from irrelevant data.

The second process of interest is that the speed or

FIGURE 1

Schematic of Broadbent-Sokolov Perceptual Information Processing Model



ease of categorizing information in either the sensory register or short-term memory is a function of both the number and availability of category states. This is equivalent to Bruner's (1957) concept of "perceptual readiness". Thus, as perceptual experience with a variety of advertising exposure situations is gained, less information of greater relevance will require processing for the categorization of events. Also, more categories could be available for such classification earlier in the perceptual process. This second process relates to the diagnosis stage of the perceptual task, and suggests a second individual difference that might be related to advertising perception.

Finally, reaction time appears to be not so much a function of the amount of incoming data, but rather of the transmission of evidence about the nature of environment. Latency is regarded as the result of the time necessary for unreliable evidence to accumulate to a point where the individual believes the probability of an error in response falls to an acceptably low level. If criteria for reaction are set to allow responses on relatively little evidence, reaction will be fast but inaccurate. It is not simple physical motor reaction speed alone which accounts for the speed of decision making.

Given the effect on decision making of the three factors Broadbent discusses, an experiment was designed to attempt to test the congruency hypothesis as well as to explore the number of category states consumers typically use in perceiving advertisements.

#### Research Design

A national probability sample of 686 English-speaking Canadian housewives completed a 67 item life style instrument. The data were analyzed to discover life style differences among users and non-users of convenience foods. Because of the product market being studied, the instrument represented items which reflected food preparation, homemaking, family life, out of home social behavior, and role-related items. An R factor analysis yielded 18 factors from the 67 items explaining 53.9% of the variance. A Q factor analysis along the 18 factors created three life style groups (an arbitrary 0.30 cut-off point on a group factor loading was used to allocate

subjects). A multiple discriminant analysis of the three groups was then performed to test the reliability of the classification system. 84% of the subjects were correctly classified in this procedure, which is a very acceptable level. The groups as constructed were viewed as potentially unique market segments for purposes of differentiating product image. The groups, labelled Specialty, Simplicity, and Basic, were comprised of subjects who had in common such R factors as:

- Specialty - enjoyment of outside activities, enjoyment of cooking and entertaining, enjoyment of dining;
- Simplicity - making cooking easier, liking prepared foods, lacking in organization;
- Basic - economy consciousness, enjoyment of family life, enjoyment of traditional meals.

The next step in preparation for the study of the congruency principle consisted of creating three print advertisements. An advertising agency was instructed to develop three communications, each congruent to the life-style groups above. The content of the product concept in life-style format was based on the four or five most heavily loaded factors which characterized each of the three market segments. Creating advertisements and/or product concepts in this manner conforms with the recommended and/or ordinary "creative" practice for the implementation of life-style research. Illustrated below in Figure 2 are the advertisements with the three life-style treatments. These advertisements illustrate how the life-style factors were operationalized by the creative people. Note the emphasis on outside activities and enjoyment of cooking and entertaining in the Specialty treatment; the simplicity and ease of preparation for the Simplicity treatment and the family and economy orientation in the Basic treatment. The most important factors characterizing that segment provide the life-style context of the product concept for each treatment.

The three treatments were then presented to 99 housewives who fell into the Specialty life-style category. These subjects were classified by submitting their 67 life-style scores from the original research instrument to



FIGURE 2

Product Concept Treatment



the original discriminant function. Each subject was asked to scale for the experiment: (1) each of the advertisements along ten dimensions (see Table 1) on a 6 point Likert type scale, and (2) to rank the advertisements in terms of their perceived effectiveness to get the subject to try the product. No time constraint was placed on this task in an effort to eliminate reaction time as a variable in the perceptual data processing task. For comparison purposes, the subjects were then exposed to all three advertisements simultaneously.

The specialty ad was held to be congruent to the subjects life-style. The simplicity advertisement was judged neutral; the basic value advertisement was judged incongruent.

Hypotheses

It can be seen that message congruency does not enter in any obvious way into the perceptual information process-

TABLE 1

## Pre-Test Attitude and Opinion Statements

## For The Advertisements

- |   |  |
|---|--|
| 1) When looking through a magazine, I would stop and read the ad.           | 6) Ranking of the potential overall effectiveness in getting you (the respondent to <u>try</u> the product.            |
| 2) The ad makes sense to people like me.                                    | 7) This ad makes me interested in the product.   |
| 3) Even if I didn't like the product, this is useful information for an ad. | 8) From the looks of this ad the people who would use the product are my type. This is the kind of ad you can believe. |
| 4) The presentation of the food is very appealing.                          | 9) The large type heading makes me interested in reading the rest of the ad.   |
| 5) The people illustrated are the ones who would use this product.          | 10) The rest of the ad tells you all you have to know about the product.   |

ing theory. Thus, the theory predicts Hypothesis 1: Specialty seeking life style consumers will not rate advertising messages designed to be congruent with their life style as being superior to messages that are incongruent or neutral.

The second process of interest relates to category states. Previous research (Lucas and Britt 1963) suggests there is, in a perceptually experienced consumer, only one category state, which seems to roughly correspond to a liking-disliking categorization. This suggests Hypothesis 2: Specialty seeking life style consumers will be found to have only one category state for the advertisements they view.

#### Analysis and Findings

Hypothesis one was tested by performing a one way hypothesis of variance on the attitude ratings for the 99 subjects taken when they were looking at all three advertisements. Table 2 presents the results. The null hypothesis of similarity is rejected in the seven dimensions (out of the ten) shown in Table 2. However, it is evident that these results do not support the conventional wisdom. The Basic food and value message is rated considerably higher on all seven of these attributes, compared to the Specialty orientad (congruent) and Simplicity directed messages. These results suggest that an incongruent advertisement would be best.

The number of attitude scales that are significantly different for these messages may be quite deceiving. On first glance it is possible to be tempted to conclude that the incongruent messages are superior along several major underlying dimensions (category states) of pre-test advertising effectiveness. Therefore, a factor analysis of the items was carried out on the ten attributes. This analysis revealed only one factor with an eigenvalue greater than one.

This eigenvalue, which had a value of 5.12, explains 51.1 percent of the variance. There is no evidence to reject Hypothesis 2.

#### Discussion and Further Analysis

The main hypothesis, that specialty seeking life style consumers would not rate specialty advertising messages as being superior, cannot be rejected. However, such consumers appear to rate a basic value advertising mes-

sage as being superior - an unexpected result. Further, the rejection of the statistical null hypothesis appears to be on the basis of a single underlying dimension or category state. It appears that consumers do not rate or rank an advertisement along several independent category states, but that an overall "halo" effect exists. The overall impression of the family-economy orientation was rated more effective in this case than the congruent "out-and-about" specialty oriented message by consumers who were specialty types in terms of their self-reported lifestyles! The notion of congruency must be questioned.

This result can certainly be rationalized, once it has been observed, in terms of the underlying theory. An incongruent or neutral advertisement represents novelty or change to the consumer, who is therefore more likely to detect it. However, further experimental work is clearly in order. One set of data is not enough.

At this stage of research it is too early to conclude that the congruency principle should never be followed. Aside from the fact that the above experiment dealt only with one product in one cultural setting, the validity of opinion and attitude ratings, and order-of-merit rankings, must be questioned (Lucas and Britt 1973). There may be a strong element of "social desirability" manifested in the ratings. The consumers studied here were not family oriented, economy conscious, nor did they care a great deal (relative to the other lifestyles) about nutritional value. As a result when exposed to what they might perceive as "more" culturally acceptable standards they may have rated the advertising along a "desirability" rather than an "effectiveness" dimension.

Another possible explanation for both the above results and the traditional congruency rule is that consumers' may use non-salient (in a choice among products sense) product attributes to adjust or create a product image so it is congruent. That is, consumers may divide product attributes into two category states: attributes that are used to make a marketplace choice and attributes that are used to create a product image congruent to the consumers' lifestyle.

When these same ads were used in another study, certain inferred product attributes were superior for the congruent advertising messages. That is, when asked to project attributes of the product concept represented by the advertisement when seen in isolation, certain perceived product attributes were rated so as to uphold the congruency principle. Table 3 is a summary table of rank correlation coefficients between advertisement

TABLE 2  
Relationships Between Mean Pre-Test Attitude Ratings  
And Message Congruency\*

Attitude scale	Specialty	Basic Food and Value	Simplicity	Significance Level
Would you stop and read the advertisement	2.39	1.99	2.29	.05
Ad has useful information	2.40	1.85	2.13	.01
Makes me interested in the product	2.85	2.26	2.52	.05
Presentation of food appealing	3.25	2.66	3.39	.01
Heading makes me interested	2.57	2.03	2.08	.01
Tells you all you have to know	2.88	1.92	2.78	.01
Would make me "try" the product	2.38	1.42	2.18	.001

TABLE 3  
Relationships Between Product Concept Congruency  
And Product Attributes (Spearman Rho)

Useful to have in cupboard .....	+ .06
Has enough ingredients .....	+ .04
Little preparation time .....	+ .22**
Not messy .....	+ .07
Substantial meal .....	- .10
Inexpensive .....	- .08
Tasty .....	+ .19*
Interesting to prepare .....	+ .38**
Healthy .....	+ .11
Different .....	+ .25**
Simple .....	+ .18
For whole family .....	+ .08
Short cooking time .....	- .08
For company .....	+ .15*
Exciting .....	+ .23**
Elegant .....	+ .20*
Good nutritional value .....	+ .25**

\* p < .05  
\*\* p < .01

congruency and perceived product attributes. Specialty was rated one, simplicity, ten, and basic value, three, in computing these statistics.

Unfortunately, there is no independent determination of which product attributes were perceived by the consumers to be salient in a choice situation. These results,

therefore, could support the image congruence hypothesis that salient product attributes should be presented in an advertisement to be congruent to the receiving consumers' lifestyle.

#### Conclusions

The congruity principle has not been supported under the conditions described here. However, when the validity of attitude and opinion ratings is reviewed, the issue of social desirability is confronted, and the fact that certain projected or inferred product attributes are rated superior for congruent messages, it is clear that much more research is required. Still, it does appear that if practitioners are going to use lifestyle in their creative advertising strategies that the congruency principle should be carefully reviewed each time in terms of the particular institutional facts known about the particular product.

#### References

Ackoff, Russell L. and James R. Emshoff, "A Reply to the Comments of Yvan Allaire", Sloan Management Review, 16 (Spring, 1975), 95-98.

Broadbent, D.E., Decision and Stress (New York, N.Y.: Academic Press, 1971).

Bruner, John S., "On Perceptual Readiness", Psychological Review, 64 (March, 1957) 123-152.

Kernan, Jerome B.; William P. D mmermuth, and Montrose S. Sommers, Promotion: An Introductory Analysis (New York, N.Y.: McGraw-Hill, 1970).

Lucas, Darrell B. and Stewart Henderson Britt, Measuring Advertising Effectiveness (New York, N.Y.: McGraw-Hill, 1963)

Mihal, William L., "Individual Differences in Perceptual-Information Processing and Their Relation To Accident Behavior", unpublished Ph.D. Dissertation, The University of Rochester, Rochester, New York: 1974.

Sokolov, E.N., "The Modeling Properties of the Nervous System", Ch. 23 in: Cole, Michael and Irving Maltzman, eds., A Handbook of Contemporary Soviet Psychology. (New York, N.Y.: Basic Books, 1969).

Tigert, Douglas J., "Are Television Audiences Really Different?", in: Fed C. Allvine, eds., Combined Proceedings 1971 Spring and Fall Conferences, American Marketing Association. Chicago, Illinois: American Marketing Association, 1972).

Wells, William D. and Douglas J. Tigert, "Activities, Interests, and Opinions", Journal of Advertising Research, 11 (August 1971) 27-35.

Whittaker, William, "The Relationship Between Individual Difference Variables, Media Usage, and Product Choice", unpublished Ph.D. Dissertation, The University of Rochester, New York: 1972.

Wright, Peter L., "On The Direct Monitoring of Cognitive Response to Advertising", pp. 220-248 in: Hughes, G. David and Michael L. Ray, eds. Buyer/Consumer Information Processing (Chapel Hill, North Carolina: University of North Carolina Press, 1974).

# CONSUMER INFORMATION ACQUISITION: PUBLIC POLICY PERSPECTIVES<sup>1</sup>

William L. Wilkie, University of Florida

## Abstract

Programs for the provision of product information to consumers are emerging as an important area in public policy. Research on information acquisition is a relatively new thrust that will provide fundamental insights relevant to the development of such programs and contribute to our understanding of consumer behavior. This paper provides an overview of this new setting and points to some significant issues inherent in conducting such research.

## Introduction

The topic of consumers' acquisition of product information is a subsector of consumer information processing (CIP) research, subject to most of its problems but also possessing interesting and important potentials. Recognizing that it is difficult or impossible to disentangle all the activities comprising consumer information processing, a useful view of this topic is that it serves to highlight a researcher's particular emphasis on an early stage of CIP. The general concern of research on information acquisition may thus be summarized as: "how much of what kinds of information consumers obtain from what sources at which times?" Although these questions have been pursued within the marketing field for some time, only recently have CIP-oriented researchers begun to address the topic. Most of the impetus for this convergence has come -- directly and indirectly -- from issues raised in public policy. The recency of these concerns is amply demonstrated by the virtual absence of coverage of this topic in the CIP "state of the art" workshop held in late 1972 (Hughes and Ray, 1974).

My thesis is that this relatively undeveloped state of research offers an opportunity for cumulative development of a research field. The pressing concerns of public policy should assist the degree of effort, while creative advances in methodology promise new forms of insight. This does not, however, alter either the incredible complexity of the CIP phenomenon nor the fact that at present we are not very far along in our understanding of same. Also, as I have argued elsewhere (Wilkie, 1975b), rarely if ever will public policy dilemmas be swiftly and surely resolved by consumer research; there are other dimensions and criteria of import which we must necessarily ignore.

The purpose of the paper is to expand upon these points. The first half provides a background and overview of emerging public policy programs for consumer information, suggesting a new context within which CIP research can be developed. The second half then turns to conceptual issues underlying the conduct of empirical research in this area.

<sup>1</sup>This paper summarizes and extends research supported by the Marketing Science Institute and the National Science Foundation (Grant GI-42037). Copies of the full report (Wilkie, 1975a) are available from the Government Printing Office.

## Public Policy Programs for Consumer Information

Readers are undoubtedly familiar with the amount and scope of recent consumer information legislation by Congress and state and local governments -- including laws concerning packaging and labeling, Truth in Lending, the banning of cigarettes from broadcast media, unit pricing, etc. In addition, executive and regulatory agencies have been quite active. The Environmental Protection Agency (EPA), for example, has undertaken its controversial gasoline mileage testing program with the intent of affording comparisons between alternative automobile brands. The Department of Transportation (DOT) is readying a program to require information on auto crashworthiness and insurability. The Consumer Product Safety Commission (CPSC) has since its inception been interested in consumer information as a means toward improving the safety experience of potentially hazardous products. The Food and Drug Administration (FDA) has required package disclosures to consumers and has instituted a program of corrective advertising for the specialist audiences receiving promotions on pharmaceuticals. The Federal Communication Commission (FCC) in the late 60's invoked its Fairness Doctrine with respect to cigarette advertising, with the intent of providing consumers with both sides of the cigarette health controversy. The Federal Trade Commission (FTC), under its broad mandate of providing fair competitive and consumer environments, has increasingly moved toward "affirmative disclosure" requirements across a number of product categories.

Currently we find more than 34 Federal agencies engaged in providing consumer information and designing consumer education programs that rely upon persuasive communication techniques. The General Services Administration (GSA) maintains the Consumer Product Information Center to centralize the dissemination of many of these publications. In sum, consumer information programs appear across a number of public policy areas, having increased substantially within the past decade.

One important characteristic evidenced in almost all past disclosure programs is an orientation toward unidimensional information provision; examples include Truth in Lending, unit pricing, care labeling, auto mileage and octane ratings. Now emerging, however is a movement toward multidimensional information in recognition of the tradeoffs between salient product characteristics in consumer choice decisions. The most likely agency for most product categories is the FTC, whose authority in this area has recently been augmented in a legal sense.

## Trade Regulation Rules (TRR's)

In essence, TRR power enables a public agency to require that certain forms of marketing practices be performed by all marketers in a generic product or service category. In the case of the Federal Trade Commission (FTC), two major court decisions and an important piece of legislation suggest that TRR programs will become a primary regulatory activity. In 1972, the Supreme Court's ruling in Sperry and Hutchinson v. FTC stressed that the Commission must undertake to develop its latent "unfairness" rationale to complement the heavily relied upon "deception" standard. This raised the possibility of actions being taken on the basis that a marketer's failure to disclose certain

product information is an "unfair" practice. In 1973, the U.S. Court of Appeals' ruling in National Petroleum Refiners Association, et. al. v. FTC strongly affirmed the agency's power to require that specified measures of gasoline octane ratings to be posted on all pumps. Most recently, in early 1975, Congress passed the Magnuson-Moss Warranty -- Federal Trade Commission Improvement Act, Title II of which specifies that FTC may prescribe Trade Regulation Rules consistent with its mandates. In addition, this law requires certain procedures to be followed in the development of a TRR.

Prominent among these procedures is the conduct of public hearings to gather pertinent viewpoints, arguments, and empirical data. There is no doubt that consumer research will be relevant for most -- if not all -- of these programs. For this reason it may be useful to briefly consider in more detail the rationale for a TRR program and the nature of policy decisions involved in its development.

There are two sorts of public policy benefits postulated to flow from the TRR program: administrative/legal and economic/social. Within the first category, it is expected that the process of developing, specifying and monitoring regulatory standards will be more efficient. For example, once the TRR has been issued the agency need only prove noncompliance to win a case; it need not consume resources attempting to prove unfairness or deception. This shift away from case-by-case litigation is also expected to improve equity for individual marketers, as the industry-wide proceeding removes the potential discrimination of being selected as a "test case," as well as providing increased certainty as to bounds on marketing efforts. Within the second category above, TRR programs for consumer information are advanced in part as a means toward approaching the "fully informed" consumer assumed to underpin the free enterprise economic system. More specifically, the availability of standardized performance information is posited to improve the efficiency of our market system by providing opportunities for consumers to compare alternatives on objective cost and benefit bases.

The degree to which these potentials will be realized, and the existence of possibly negative outcomes, obviously depend on the quality of each TRR's coverage and specifications. Inputs from consumer researchers will be sought and relied upon during the process of a TRR's specifications; a brief overview of such program decisions may be helpful in placing closer perspectives on the multitude of issues likely to arise. In considering the following discussion, however, it is particularly important that we recognize that this process is susceptible to numerous influences (philosophic, legal, economic, and political) apart from strictly technical concerns.

#### TRR Program Decisions

While it may at first blush appear that design of an effective consumer information program would be quite straightforward (i.e., simply find out what information consumers need, then require that it be made available), in actuality these programs will be extremely difficult to develop. The roots of difficulty relate to both non-consumer matters (e.g., the policy influences noted above, plus possible technical problems with standards and test methods) and to CIP research concerns. For example, do consumers currently understand the generic product in question (i.e., do they know what information they need?) and do they possess the necessary cognitive means to process or otherwise use such information?

Given these sorts of issues, policymakers face a more involved, six-stage, decision process (Day, 1975; Wilkie, 1975a). Decisions made at each stage must take into account problems likely to be encountered at subsequent stages, thus introducing simultaneity as a factor.

The first stage is the selection of topic (i.e., TRR program), which involves numerous policy considerations and may well be the most subjective area of our sequence. With regard to information, moreover, there are two alternative approaches that might be taken. The first, more typical to date, would stress one particular dimension of information which is applicable across a range of product and service categories. Examples of such programs include cost of credit (Truth in Lending), nutritional disclosures, and, in the near future, energy consumption in product operation. The second approach is quite new, but is likely to receive considerable future attention. Here the focus is on only a single product or service class, but with options of covering a number of information dimensions pertinent to that class. Because this approach raises particularly fundamental CIP research issues, it will be adopted in the remainder of this paper's discussion.

Criteria used for the selection of information topics are beyond the scope of this paper; they can reduce to the view that the value of information in a given category is a function of the expected opportunity losses from suboptimal choices in that category. It presently appears that three general product types will receive primary attention: those in which safety is an issue, those pertaining to health (especially food and drugs) and those with considerable economic potentials such as consumer durables.

In addition to product selection, the matter of appropriate program objectives is a critical question deserving careful attention, as these objectives will directly affect the nature of all subsequent program decisions as well as determining relevant CIP research designs and criteria (dependent variables). The basic question, of course, is "what kinds of effects are desired at the consumer level?" A fundamental policy issue, however, involves the extent to which government may act so as to direct or influence free consumer choice in the marketplace. It may be generally agreed that information from a government program should meet the dual criteria of being both neutral (in the nondiscriminatory sense) and effective in the sense of being useful to consumers (Wilkie and Gardner, 1974). There is little doubt that the presumed neutrality on brand choice (i.e., consumers can be left to make their own choices once the needed information has been made available to them) is a major appeal of these programs.

The choice of program objectives will play an important role throughout the remaining program and decision stages, helping to direct not only the role of consumer research, but also providing general criteria for how much, which kind, and the exact manners in which information will be required to be made available to consumers. Following a brief presentation of each of the remaining decision stages, we will return to the question of appropriate objectives.

The second stage, identification of information dimensions, will likely stress objective and/or performance characteristics of the product, as these are most susceptible to standardization. There are two basic sources for the determination of relevant product characteristics: technical expertise and marketing or consumer research expertise. Day (1975) reports that most technical experts responding to a 1972 FTC task force study on this question agreed that significant product

characteristics could be identified from industry knowledge and consumer surveys. It is possible, however, that consumers, who have not previously been exposed to such information would have difficulty reporting a need or desire for same. If it happens that experts and consumers don't agree, policymakers will confront the question of whether a proposed program is creating needs rather than responding to them--an issue underlying many of the decisions to be made.

The third stage, development of test standards and measures, moves to engineering test protocols to yield standardized ratings for each brand on each characteristic. Reliance will be placed on technical experts, perhaps involving the National Bureau of Standards and various industry standard-setting bodies. Day's (1975) analysis of the task force study reports a general consensus that suitable test methods for durable goods either presently exist or can be developed. There were noted, however, several areas of concern which one might expect to become significant once a given program is underway. For example, the high test cost associated with some sophisticated products may well be disproportionately burdensome to smaller competitors. Because various materials perform differently as a function of test method, moreover, difficulty in obtaining industry-wide agreement on a single method can be anticipated. Also, some products would have to be measured in terms of their effects on individual consumers which could necessitate dealing with idiosyncratic reactions as a function of individual differences. In sum, although technical experts were generally optimistic, we should not expect any given program to proceed smoothly through this stage.

At this point the raw materials for the TRR program are available, and attention turns to communication decisions. The fourth stage, determination of reporting format, involves three primary issues: (1) how many and which information dimensions should be included? (2) in what form should the information appear? and (3) what problems, if any, are encountered if some difficult-to-rate characteristics (e.g., convenience, service, credit terms, guarantees) are not included?

Day (1975) relates that the 1972 FTC task force report originally proposed that only two or three characteristics will usually need to be measured for a product class, and that these could be combined (i.e., weighted and summed) into a single, overall, "grade" to be reported to the consumer. Both suggestions were strongly criticized by respondents to the proposal, the first on the basis that two or three characteristics are simply not enough to constitute a meaningful basis for determining or explaining differences in performance. The weighting model also encountered several legitimate criticisms, especially that it incorporates the assumptions that all consumers assign values to performance dimensions (i.e., they trade-off benefits) in the same way, and that either all performance characteristics possess an identifiable "ideal" level or that more is always better. A further problem with reporting a single grade is that consumers will be unable to ascertain why one brand might be graded lower than another, thus rendering trade-offs between measured and unmeasured attributes impossible.

The general complexity of program decisions is by this point becoming apparent. At one extreme, policymakers could provide (require) all the information generated by technical experts, maintaining all dimensions and reporting all results in technically precise terms. At the other extreme, only a summary grade might be used. The primary difficulties of these policy decisions spring from anticipated constraints of consumer infor-

mation processing, and involved issues of "meaning", simplification, efficiency, and equity.

Closely related to these questions are those encountered in selecting dissemination channels (stage five). Included here are questions of sources, channels, and points of consumer exposure. Unlike the previous stage's difficulties imposed by the proper role of the government actor, this stage may well benefit from such a role.

FTC is understandably reluctant to centralize the dissemination process, which indicates that marketers will probably be required to make the information disclosures using some combination of their normal channels. Possibilities include package labels, product information tags, point-of-sale materials, verbal explanation by sales persons, and/or advertising vehicles.

From the consumer's viewpoint, he therefore could have the information available to him when he is not contemplating the purchase, during pre-purchase planning activities, as well as at point of purchase. Allied to this dissemination -- although not within FTC's control -- are consumer education programs that might be aimed at developing consumers' CIP abilities, private product rating publications, and/or voluntary programs by brand marketers.

A broader question involving both format and channels concerns the extent to which brand comparisons by the consumer can or should be fostered, and where in his decision process these comparisons would occur. Some alternatives (e.g., dimensional disclosure in broadcast advertising) could place heavy stress on consumer memory. Other alternatives (e.g., summary ratings at point of sale) could hinder comparisons with brands not available. It seems quite likely, therefore, that combinations of formats and channels must be utilized.

Our final decision stage involves structuring a formal evaluation and program revision process, and reflects two major issues -- the fact of a formal evaluation and the importance of clearly defined objectives. Program evaluation -- especially of consumer impacts -- has been too often ignored by public agencies. Three benefits are offered by planning for evaluation as an integral decision stage:

1. "Before" measures of the current state of the market and consumer environment can provide useful input to the five decision stages above, in addition to serving as baseline measures;
2. "After" measures reveal not only the actual impacts of the program, but can also be used as guides to revise or improve the ongoing efforts;
3. In the broader sense, formal evaluations provide "learning feedback" for those working in the public policy field, this being most valuable in the design of future programs. It is apparent that this hasn't happened from most of our past programs--as a society, we have not advanced much in our understanding of public programs for consumer information provision.

The observed lack of formal evaluations probably does not reflect disinterest in the program or policymakers' desires to remain ignorant of its effects. Rather, it would appear to spring from two other sources: (1) barriers to conducting objective evaluation research in an action setting, and (2) serious difficulties in specifying program goals or objectives. The first of these is beyond the scope of this paper, dealing with questions such as the research setting within the

organization and program, perceived needs for actionable conclusions, and the general negative cast of evaluation research (Weiss, 1972; Wilkie, 1975a).

The second point -- difficulties in determining appropriate and precise objectives -- is a critical one for consumer researchers, as these provide the bases for meaningful selection of dependent variables. A primary issue for almost all research in this area is the proper weighting of consumer "cognition" versus changes and choice behavior. Consider, for example, decisions on information format (stage 4). A stress on consumer cognition (including product understanding as well as brand knowledge) would assume that consumers want and can use more product information. Emphasis would be on "full" disclosure with a high degree of specificity; overall grades would not be used, and the number of information dimensions might easily be high. Such a disclosure would be expected to require increased cognitive effort from the consumer. Brand alternatives are more fully described, therefore subject to more combinations of trade-off and perhaps decreased reliance on overall evaluations of brand superiority and inferiority. Thus we might also expect to see more alternatives under active consideration. Two negative effects might also, however, accompany this stress. Some consumers, anticipating that they cannot (or are not willing to) handle the increased complexity, might ignore the information. Others, in attempting to process it, could encounter increased anxiety about the decision and increased uncertainty in their ability to make a choice.

A stress on consumer choice behavior, conversely, would place more emphasis on resultant decisions than on the information per se. Policy decisions would be quite complex with this approach, in that one would have to consider models of consumer choice, make some resolutions between technical dimensions and consumer perceptual dimensions, and deal with problems relating to dimensions excluded from program disclosure, particularly those that are hard to measure but do relate to product performance. "Simplification," or provision of summary forms of ratings will need to be considered here, promising more efficient consumer processing and a likelihood of greater utilization. Against these potential benefits, however, policymakers must balance the tendency toward paternalism (especially if there is any question that products would be categorized differently if dimensions were weighted differently), and the charge of withholding certain information.

In summary, the public policy environment is fraught with nuance; hard answers will simply not be available for CIP researchers to build studies upon. This does not suggest that empirical studies aren't needed, but does mean that we researchers must carefully consider the problem setting when planning the study. Also, of course, it suggests that without such consideration, what appear to be direct policy implications from our typical research will rarely turn out to be nearly so meaningful as the researcher believes them to be.

#### Research Orientation

In an earlier assessment of the CIP literature (Wilkie, 1975a) I have argued that surprisingly little work which might accurately be labeled "consumer information processing" research has been undertaken. The vast majority of relevant work on "information processing" is found in nonconsumer fields, using different sorts of "information" as stimuli, often within the context of complex problem-solving (e.g., Berlyne,

1965; Newell and Simon, 1972). Most consumer research, moreover, has been conducted within the framework of marketing's controllable variables, especially persuasive communication. Here the stress has typically been on the summary effects of a stimulus rather than on intermediate "processing" activities; neither designs nor measures have been oriented to the study of processes.

Emerging research on consumers' acquisition of information is moving away from these boundaries -- an encouraging development. We are, however, still at an early point in such research. Recognition of some important distinctions in orientation may therefore be particularly useful for further progress in this field.

There are essentially three characteristics of the public policy orientation which differ substantially in nature from our prevailing consumer research orientation:

- (1) neutrality on brand choice,
- (2) a changed consumer environment, and
- (3) little direct interest in source effects.

Most basic is the issue of neutrality with respect to brand choices of consumers. There are three aspects to this orientation which are highly significant. First, "persuasion" is not at issue for public programs (apart from attempts to insure that the information itself can be accessed and utilized by consumers). In terms of utilization, moreover, public policy programs are not directional in nature. This is obviously in contrast to the overtly directional goals of marketing and advertising (i.e., to influence each target consumer to be predisposed to purchase a particular brand). Second, in keeping with this overriding constraint, public policy must deal only with "objective" or factual information of a standardized sort, typically dealing with specific dimensions of product makeup or functional performance. Subjective appeals or dimensions ("brand image", aesthetic, styling, etc.) appear not to be possible for inclusion in these programs, though they are major factors for marketing and advertising programs. Third, and following the above points, the underlying CIP model applicable to these programs must be cognitive in nature.

Another class of considerations characterizing public policy programs involves their necessary concern with a sharply altered or changed consumer information environment. The essence of such change is fostered by the program itself, presenting what may be entirely new dimensions of information to consumers, and presenting brand data for each competitive offering on each dimension. The possible CIP implications of an altered environment are substantial; serious consideration must be given to new forms of research design, measurement, and criterion variables.

A third class of issues involves the relative importance of the various elements of CIP research concerns, in that the priorities of policy interest in CIP will differ. A major area of difference, for example, involves emphasis on "source effects" such as selective attention or source credibility. Given that the "information" of public policy is intended to be objective, standardized, and neutral, it should not differ by individual in terms of its perceived truth, but only by its value or salience within CIP and choice processes. In keeping with this point, there may well be little direct interest evidenced in the impacts of specific marketing vehicles such as advertising, especially in terms of their non-cognitive operations.



## Issues For Empirical Research

In addition to shifts in orientation to account for the public policy setting, researchers must also face the added complexity of the topic of consumer information processing. As noted earlier, we do not have a rich body of concepts or findings currently available on this topic. Drawing from work in a number of related disciplines, however, it is possible to derive a few basic generalizations likely to apply to CIP (Wilkie, 1975a):

1. Variability by Adaptation. Past research has found that the processes invoked by humans appear to heavily depend upon (a) the task they are faced with and (b) the context within which that task is undertaken. At its start, then, information processing appears to be adaptive in nature. This suggests that CIP researchers must take particular care in the design of the research setting, in mitigating demand characteristics, and especially in generalizing from a study's results.
2. Importance of Memory. Processing occurs "between" stimulus and long-term memory, thus is subject to two flows rather than only one. Study of the latter is at least as important as study of the former.
3. Capacity Constraints. "Buffers" are presumed to exist for the purposes of short-term retention and symbol manipulation. These buffers have severe capacity (size and temporal) limitations, leading to limitations on consumers' ability to process information. Consumers are not computers, they will not automatically process everything they are exposed to.
4. Subproblem Processing. These limitations lead to non-optimizing subproblem processing for relatively complex tasks. The problem is broken into smaller, more manageable sub-decisions, with information relevant to each of these being selectively manipulated. For CIP research, this point suggests that we need to explicitly view consumers as proceeding in relatively discrete stages, and try to measure results at each stage. Recent research methods offer considerable potential to improve our understanding here.

The net results of these "generalizations" suggest that considerable activity goes on between stimulus and response, and that investigation of these issues provides problems and opportunities for CIP researchers.

It is evident that any single CIP study cannot hope to address all the aspects of the phenomenon. At the same time, however, researchers must be careful to attend to the possible impacts that some of these aspects may have on results and conclusions. In particular, there are five classes of underlying issues that will apply to most or all of research in this area:

1. Motivational Factors
2. System Capacities
3. Existing Consumer Knowledge and Predispositions
4. System Invariance
5. Individual Differences

Our brief coverage of these zones should properly be viewed as speculative, since virtually none of the underlying research has been conducted within the consumer context.

### Motivational Factors

Consumer information processing is dynamic; it occurs

through time. The role of motivational factors is thus to explain the "energy" dimension of CIP: how is it started and how is it sustained? Public policy's interest here relates to consumers' desire to acquire and process the new information available to them. Relevant literatures include comprehensive consumer behavior models, activation, arousal, drives, cognitive consistency, directed thinking, variety-seeking, and perceived risk.

Of particular import for empirical research on consumer information acquisition is the likelihood that respondents' motivations within a study will differ substantially from those occurring in the naturalistic environment. Not only may the energy level be heightened, but -- of equal concern -- certain forms of task goals may be applied. This suggests that severe problems may be encountered in attempting to describe naturalistic information acquisition. Included in this sector is the question of gaining information when not engaged in problem-solving. Processes of such acquisitions are also fair game for investigation.

### System Capacities

The issue of system capacity refers to the consistent finding that humans are subject to finite limits to assimilate and process information during a given time period. The introduction of numerous information dimensions might therefore overload rather than assist acquisition and processing. However, capacity constraints do not apply to all CIP sectors; sensory receptors and long-term memory (LTM) possess vast capacities. This suggests further investigations into consumers' employment of LTM to reduce burdens on short-term memory and storage. Possibilities include fostering CIP "rules," provision for external "writing" and/or reaccessing of information, assisting sequential comparisons, development of information "chunks" (see Miller, 1956; Simon, 1974), and lessening the precision of task-required judgments. Also -- of considerable import for external validity in the public policy consumer setting--the time constraints employed in task studies must in some manner be altered to correspond to the time and learning freedoms of the naturalistic consumer world.

This topic has already received some attention by CIP researchers, notably in the studies by Jacoby and associates (Jacoby, Speller, and Kohn, 1974a, 1974b). Further perspectives on research issues here are available in Russo (1974), Summers (1974), Wilkie (1974), and Jacoby (1975).

### Existing Knowledge and Predispositions

This third "problem zone" concerns the aforementioned sector of a consumer's long-term memory (LTM). In addition to playing a major role in the nature of CIP when a consumer is exposed to an informational stimulus, this zone affords researchers their primary opportunities for measuring the effects of information. There are three central concepts of interest in this zone: (1) consumers' conceptual structures, or rules used to guide information processing, (2) cognitions of (i.e., knowledge about) the product class and alternatives within it, and (3) predispositions, including affect and preferences, for those alternative brands. Much recent consumer research attention has been devoted to consumer predispositions, while very little work has focused on consumer cognitions and/or conceptual structures.

Research on "multi-attribute attitude models" has especially been aimed at assessing the nature of brand predispositions (see Green and Wind, 1973; Wilkie and

Pessemier, 1973; and Cohen, 1974 for overviews of this research stream). Attitude models have, however, focused on static measurements rather than active CIP, such that their treatment of "attributes" and "ratings" differs from the sense of these terms in the public policy setting.

Attributes, in our general framework for CIP, represent objective content or performance characteristics of a given product. Attitude models, in contrast, generally define attributes in terms of consumers' subjective (i.e., perceptual) dimensions. Because of the role of "organization" in LTM, we do not expect a one-to-one correspondence between such objectively and subjectively-defined dimensions. Two implications derived from this distinction. First, this may in part account for the rather consistent findings (in attitude research) that rather few -- about five -- attributes are sufficient to account for brand predispositions. Second, the extent of non-correspondence between the two kinds of dimensions should be assessed in terms of possible problems in translating the objective information to the subjective dimensions in LTM.

"Ratings" likewise differ in the two contexts. Our CIP framework treats ratings as the objective amount provided by a given brand on a given attribute. Attitude models, conversely, treat them as "beliefs" or "expectancies", reflecting each consumer's subjective perception of how well a brand performs with respect to an attribute. It is especially important that researchers recognize this distinction, as research on the attitude model has in the past significantly relied upon differences between consumers in their beliefs (ratings) concerning brand/attribute values. The public policy CIP context clearly suggests no differences between consumers on the ratings per se. As with attributes, the research implication of this distinction indicates that CIP research stress should be given to questions of how a consumer's existing perceptions will effect his interpretation and integration of new information.

Measurement of a consumer's prior brand information represents a related approach which is less developed in the consumer research literature. As discussed by Woodruff (1972) this approach attempts to assess a consumer's cognitions concerning the actual ratings of brands on specified product attributes, and uses obtained data as measures of the task information on that parameter. Based upon normative concepts of Bayesian decision theory, this approach attends directly to issues of "uncertainty" in existing consumer knowledge and the role of external information as it serves to alter such uncertainty. The methodologies and perspectives of this research stream appears to offer considerable potential for future CIP research application, especially regarding the incorporation of uncertainty into models and measures.

#### System Invariance

While the above zone concerned issues of validity in measuring static structures underlying CIP, the problem zone of "system invariance" extends to issues of both fixity and change in such structures. There are two areas of concern in this regard: (1) can CIP structures be accurately measured in a static manner? and (2) how can true change in such structures be assessed? The first concern in this zone -- termed "fixity" -- involves several subissues, including test-retest reliability, differences due to use contexts, and the rule of uncertainty in existing conceptual structures.

The second concern -- true change in conceptual structure -- relates to the possibility of consumer "learning" as a result of exposure to new product information. Learning in this sense can refer, not only to the acquisition of content (i.e., brand ratings), but also to increased awareness of additional brands and product attributes, as well as a consumer's development of new relational rules. We should recognize that a chronic problem with studies conducted at a single point in time is their inability to account for such learning, thus very likely understating the potentials of a given informational stimulus over time (e.g. public policy programs or advertising campaigns) to effect changes in consumers' knowledge and behaviors.

#### Individual Differences

The fact that consumers differ raises difficult problems for both strategists attempting to design informational stimuli and for researchers attempting to assess their effects on CIP. The most notable results of these problems are real difficulties in arriving at meaningful "generalizations" and in developing programs aimed at "the" consumer. Also, due to individual differences in values and utilities, researchers are often unable even to work with precise levels of criterion variables. "Choice quality" is a good example; we cannot necessarily conclude that one brand is a better choice than another, especially if prices differ. Considerable extra effort is required (e.g., Jacoby, Speller, and Kohn, 1974; Summers, 1974) simply to obtain a single dependent variable. Other difficulties associated with this zone include: a tendency to employ cross-sectional analyses (e.g., correlation, regression) which assume respondent homogeneity, a cost of "trait" focus in tending to decrease attention to processing activities, and a need for larger sample sizes. One promising development in this regard, however, is the progress of Bettman's attempts (Bettman, 1974; Bettman and Jacoby, 1975) to develop a statistics of observed processes.

#### Conclusion

The discussions in this paper have briefly posed a series of contextual and conceptual issues for CIP research. It has not dealt particularly with research on information acquisition, primarily because this subfield of CIP must also grapple with these basic concerns. As noted at the start of this paper, this field is still new and at an active stage of development. Several recent studies are indicative of the promise of this research, including the creative research paradigms of Jacoby and his students (e.g., Jacoby, 1975) methodological advances of Bettman, and the conceptual work of Russo (1975). The public policy setting has sparked much of this interest, but brings with it a need to shift away from some fairly entrenched consumer research orientations. It is hoped that this paper will assist in the recognition of such distinctions and the development of a major field of research.

#### References

- Daniel E. Berlyne, Structure and Direction in Thinking (New York: John Wiley & Sons, Inc., 1965).
- James R. Bettman, "Toward a 'Statistics' of Consumer Decision Net Models," Journal of Consumer Research, 1 (June, 1974).
- James R. Bettman and Jacob Jacoby, "Patterns of Processing in Consumer Information Acquisition," in B.

Anderson (ed.), Advances in Consumer Research: Volume 3 (Cincinnati: Association for Consumer Research, 1975).

Joel B. Cohen, "Toward an Integrated Use of Expectancy-Value Attitude Models," in G. B. Hughes and M. L. Ray (eds.) Buyer/Consumer Information Processing (Chapel Hill, N.C.: University of North Carolina Press, 1974).

George S. Day, "Full Disclosure of Comparative Performance Information to Consumers: Problems and Prospects," Journal of Contemporary Business, 4, 1 (January, 1975).

Paul Green and Yoram Wind, Multiattribute Decisions in Marketing (New York: Holt, Rinehart, and Winston, 1973).

G. David Hughes and Michael L. Ray, Buyer/Consumer Information Processing (Chapel Hill, N.C.: University of North Carolina Press, 1974).

Jacob Jacoby, "Perspectives on a Consumer Information Processing Research Program," Communication Research, 2, 3 (July, 1975).

Jacob Jacoby, Donald Speller and Carol A. Kohn, "Brand Choice Behavior as a Function of Information Load," Journal of Marketing Research, 11 (February, 1974a), pp. 63-9.

Jacob Jacoby, Donald Speller and Carol A. Kohn, "Brand Choice Behavior as a Function of Information Load: Replication and Extension," Journal of Consumer Research, 1 (June, 1974b).

George A. Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," Psychological Review, 63, (March, 1956), pp. 81-97.

Allen Newell and Herbert A. Simon, Human Problem Solving (Englewood Cliffs, N.J.: Prentice-Hall, 1972).

J. Edward Russo, "More Information is Better: A Re-evaluation of Jacoby, Speller, and Kohn," Journal of Consumer Research, 1, 4 (December, 1974).

J. Edward Russo, "An Information Processing Analysis of Point-of Purchase Decisions," (paper presented at the Fall Educators' Conference, American Marketing Association, Rochester, New York, August, 1975).

Herbert A. Simon, "How Big is a Chunk?" Science, 183 (February 8, 1974), pp. 482-88.

John O. Summers, "Less Information is Better?" Journal of Marketing Research, 11 (November, 1974).

Carol H. Weiss, Evaluation Research (Englewood Cliffs, N.J.: Prentice-Hall, 1972).

William L. Wilkie, How Consumers Use Product Information (Washington, D.C.: Government Printing Office, 1975a).

William L. Wilkie, "Consumer Information Processing Research: Product Labeling," (working paper series, Marketing Science Institute March, 1975b).

William L. Wilkie, "Analysis of Effects of Information Load," Journal of Marketing Research, 11 (November, 1974).

William L. Wilkie and David Gardner, "The Role of Marketing Research in Public Policy Decisionmaking," Journal of Marketing, 38 (January, 1974), pp. 38-47.

William L. Wilkie and Edgar A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10 (November, 1973), pp. 428-41.

Robert B. Woodruff, "Measurement of Consumer's Prior Brand Information," Journal of Marketing Research, 9 (August, 1972), pp. 258-263.

# THE IMPACT OF MOBILITY AND SOCIAL INTEGRATION ON INFORMATION SEEKING

Donald J. Hempel, University of Connecticut  
William J. McEwen, University of Connecticut

## Abstract

This paper integrates considerations of some social and psychological factors which may determine information utilization patterns in consumer decision making. A theoretical framework is provided and cross-cultural comparisons drawn from two studies of housing purchase are presented as partial indication of the role of two major factors affecting use of and preference for alternative information sources.

## Introduction

Consumers in all societies are experiencing the well publicized problems of dealing with an ever-widening information explosion. This is a consequence of expanding information requirements (for each choice situation) as well as of a rapidly growing variety and complexity of consumption decisions. A major concern of planners must therefore be the effective channeling of communications to satisfy these increasing consumer needs. Studies of similarities and differences in information utilization which examine patterns across social systems (e.g., regional and national boundaries), however, have been confined largely to descriptive statements embodied in the diffusion of innovations literature.

While considerable research has addressed the psychological determinants and correlates of information processing, such literature seems inadequate to the task of defining inter-community utilization differences. Of more pragmatic importance to the communications planner attempting to facilitate search and utilization in various sub-populations (e.g., mover segments) would be a more aggregative examination of systemic factors influencing information "handling." Thus, extending the largely descriptive innovations approach via a focus on the interplay of both psychological and social factors influencing information utilization in consumer decision making seems particularly useful.

One of the most consequential of all consumer decisions involves the choice of family residence. The changing spatial location patterns in a society reflect fundamental adjustments in the labor force and housing stock, as well as the adaptability and propensity to change of different population segments. The various aspects of spatial mobility can also serve as useful indicants of social system development. While the functional aspects of mobility have been studied extensively, the dysfunctional consequences still remain largely unexplored in empirical terms. This is particularly characteristic of knowledge concerning the impact of changes in physical environment on information needs and usage behavior.

These are several basic reasons why this problem area is of considerable theoretical and practical significance:

1. Those experiencing high rates of physical mobility may have special difficulties in adjusting and responding to new social environments (e.g., Toffler, 1970; Packard, 1972).

2. Those who came from different regional or national social systems bring varied bases of experience and have other characteristics which may enhance their effectiveness as transmitters of new ideas.
3. Physical mobility is important as a life style attribute which can be readily measured and effectively used as a basis for market segmentation.

## The Conceptual Framework

The conceptual scheme presented here attempts to develop a broad framework for analyzing the interaction of social and psychological variables in consumer information utilization behavior. This represents a literature review plus implications derived from several recent studies of social systems conducted by the authors. The data base used in the following section serves primarily to indicate how this scheme may be in part operationalized. While a complete application of the conceptual framework is not feasible with the data at hand, work is now underway to extend this analysis into a study of related social systems.

## Physical Mobility

A review of the available literature suggests a basic distinction between two types of physical mobility. The first type involves relatively long durational expectancies and thus requires rather substantial effort to adapt to norms of a new social system. This "environmental commitment" is especially characteristic of residential changes, and, to a lesser degree, of occupational changes. This contrasts with a second type involving short durational expectancies with little perceived need for adaptation and easy reversibility of behavior patterns. The "environmental transience" reflected in this classification is represented by various forms of travel, ranging from short pleasure or business trips to extended vacations. This study is concerned with the impact of both types of mobility upon the information-seeking behavior of families involved in changing their residential environment.

Discussion of the consequences of long durational changes in residence or occupation shows disagreement among social observers. Packard (1972) has maintained that residential mobility--increasingly involving changes across community social systems--has social alienation consequences. He has maintained that spatial mobility (of the environmental commitment type) has rather inherent dysfunctional consequences, even for upward movers. In contrast, Gans (1972) has maintained, in line with much of the available organizational research (e.g., Vroom, 1969) that social isolation seems more frequently a result of non-voluntary spatial mobility (e.g., urban renewal), especially when the cross-system changes involve major life style alterations. It thus appears that the dysfunctional consequences of spatial mobility involving environmental commitment will depend, at least in part, on similarities in what Sheth and Sethi (1973) have termed the "cultural life style." Similarity in cultural life style across systems should ease the potential strain resultant from voluntary, or even non-voluntary, shifts in residence and/or occupation. Thus, the move from a large city in one country

to a large city in another country may require less cultural adaptation than the change from a rural to an urban area within the same country.

#### Mobility and Search Behavior

One major area of study concerning possible mobility-induced consequences has centered on the notion of decision making and the nature of information usage. Although seldom studied directly, a part of the concern of social observers such as Packard (1972) and Toffler (1970) has to do with information "coping" behavior. As mentioned before, Packard's thesis is that mobility hinders general coping, stemming, in part, from the fact that normal channels of information search and sharing are disrupted. That there are differences in normative information usage patterns across cultural and even subcultural social systems has been established (e.g., Dervin and Greenberg, 1972). Where cultural life style differs, information-related behaviors (as one set of activities comprising this life style variable) also differ. The question centers on the extent to which different types of information search patterns are associated with decision processes which have been affected by physical mobility.

While some differences exist as to the ultimate consequences of long term spatial mobility, there has been apparent agreement that short term environmental transience is functional both for the individual (as a quality of life correlate and as one indication of openness to innovation) and for the social system (as a general correlate of system development). This may simply reflect the fact that this type of spatial mobility, in contrast with the first, almost always involves voluntary changes which are rather readily reversible and hence imply little commitment.

Rogers (1969; 1971) has proposed that certain types of short term mobility facilitate information search. Thus it would be anticipated that those who exhibit higher environmental transience mobility ("cosmopolitaness") should have different usage patterns than those low in this aspect of mobility. Rogers presents some evidence that the former should utilize more deliberate and efficient search patterns. Whether this difference arises as a consequence of the increased external contact provided by environmental transience mobility or, perhaps more probably, as a function of personality and motivational differences is actually a separate issue. The precise causative factor may not be as crucial to the information planner as the fact (and nature) of differences in information search patterns.

#### Social Integration

The adaptation and consequent alteration in information search patterns which are required with environmental commitment mobility are not manifested equally in all consumers. Adjustment to a new communication environment is affected by the similarity of prior cultural life style and by the extent of previously experienced physical mobility--of both the commitment and transient variety. One further variable of importance involves the consumer's level of social integration. Extent of involvement in the social system reflects both adherence to system norms and frequency of (and reliance on) intra-system contacts (Rogers, 1971). This may involve an overall higher level of communication activity with greater usage of system information sources (especially other people). As such, this would thus appear to reflect a general lack of felt "alienation" and consequently, as McLeod

et al. (1966) have suggested, a greater reliance on certain types of communication sources (e.g., in-depth print media treatments).

#### Social Integration and Search Behavior

Packard (1972) has also suggested the potential for communication-related involvement consequences. He has proposed that social integration can act to lessen certain of the negative aspects associated with residential mobility. It is suggested that, for example, belonging to a social organization will reduce the sense of rootlessness resultant from required changes in environment, if the organization (or one closely parallel) has chapters or member groups in that new social system. Without this transference opportunity, however, even greater strain may be placed on those moving. What is proposed, then, is that contact with continuing social organizations may provide an extension of the former social system--or at least reinforce a familiar set of already shared norms. This sense of continuity or familiarity amounts to a reduction in cultural life style dissimilarity.

Furthermore, social integration seems to modify information channel exposure and influence communication opportunities. The common organization may, by means of the contacts made through these groups, facilitate information search. The result should be a group-induced and/or supported difference in information source utilization, most notable with regard to major decisions such as are involved in family housing choice.

Ultimately, it is the combination of mobility and social integration which is likely to determine how one adjusts to a new residential environment. The interaction of these factors operates both to define information needs and to provide opportunities for obtaining information. This adjustment process has important implications for overall satisfaction with life, as well as for many subsequent consumption decisions (Andreasen, 1966). The first impressions of one's living (residential) environment formed during the period of housing search contribute significantly to the perceptual framework which is then used in the selection of shopping facilities, churches, commuting routes, and other "institutional connections." These choices, in turn, influence the buyer's further exposure to information channels and to the specific sources which are used in later purchase decisions. Thus, the information channels used by families involved in residential change--especially where the degree of change is substantial--can be studied as a generative system with far-reaching implications for communication planners.

#### Application Of The Conceptual Framework

##### Data Base

The data for this study were obtained from two surveys of recent home buyers conducted during the summers of 1968 and 1971. 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. Personal interviews and mail questionnaires were used to obtain information from both husbands and wives in 206 households from the Hartford metropolitan area of Connecticut, and 317 households in the Preston-Lancaster area of Northwestern England. The bases for analysis of the mail-questionnaire data presented here was reduced by nonreturns and the elimination of observations with missing information. A response rate of almost 70%

was obtained for both samples. Details concerning the research design and questionnaires have been published elsewhere (Hempel, 1970).

#### Measurement

The communication channels used in most purchase decision processes can be classified into two major categories: consumer-dominated and marketer-dominated (Cox, 1961). These two channels may function as complementary information disseminators with considerable variance in their effects upon purchase behavior. The relative importance of different information sources used by the home buyer has been shown to vary with the stage of the decision process and a number of consumer attributes (Hempel, 1969).

In order to examine the general patterns of information seeking among home buyers--in terms of both specific source usage and cumulative measures of channel usage--three aspects of information seeking behavior can be analyzed: (1) exposure--what proportion of the buyers recall using the source while looking for and purchasing their home; (2) sequence--among those using the source, what was the order in which each source was first contacted; and (3) evaluation--which sources were recommended to other buyers. These measures of source importance were obtained for each of 16 sources selected to represent four different types of general information channels--interpersonal communications, mass media, commercial sellers, and self-initiative.

TABLE 1  
Patterns of Information Source Utilization in  
Connecticut (N=206) and Northwest England (N=292)

Information Source	Measure of Information Source Utilization					
	Exposure - % Using Source		Sequence-Median Order of Contact <sup>a</sup>		Evaluation - % of Exposed Recommending to Other Buyers	
	HTC	NWE	HTC	NWE	HTC	NWE
<b>A. Personal</b>						
1. Friends	51%	44%	3.0	3.0	20%	9%
2. Co-workers	28	19	3.0	3.0	4	2
3. Relatives	27	34	3.0	3.0	7	7
4. Residents of Neighborhood	16	16	5.0	5.0	6	4
Composite Index (PSX)	72	61	3.0	2.5	28	16
<b>B. Mass Media</b>						
1. Newspaper Ads	88%	79%	1.0	1.0	34%	27%
2. Newspaper Articles	26	10	3.0	3.0	5	2
3. Magazines or Pamphlets	11	3	3.0	3.0	2	2
4. Books	5	1	3.0	3.0	1	0
Composite Index (MMX)	89	79	1.5	1.0	35	30
<b>C. Formal Commercial</b>						
1. Real Estate Firms	83%	74%	3.0	3.0	57%	38%
2. Builders	24	16	5.0	4.0	5	6
3. Employees of Lending Institutions	20	36	5.0	5.0	13	17
4. Attorneys	27	56	5.0	4.0	17	27
Composite Index (FCX)	94	81	3.0	3.0	72	71
<b>D. Self-Initiative</b>						
1. Owner of Property	16%	16%	5.0	4.0	2%	1%
2. Walking/Riding Around	55	65	4.0	3.0	11	13
3. For Sale Signs	27	48	4.0	4.0	2	2
4. Town Records	10	4	5.0	6.0	2	2
Composite Index (SIX)	63	70	4.0	3.0	17	18

<sup>a</sup>The median order of contact (sequence) was calculated on the basis of those who were exposed to the source. The actual number of observations for these estimates can be determined by multiplying the subsample size (N) by the percentage shown for the exposure measure.

An index of channel usage was constructed for each aspect of information seeking to reflect the cumulative effects of the four sources in each channel. The exposure and evaluation indexes reflect whether or not the household referred to one or more of the four specific sources within each set. The sequence index represents the mean order in which the specific subset of sources used were first contacted.

The functioning and usage of the information channels as measured by these four indexes were studied in terms of their covariance with physical mobility and social integration. Environmental commitment mobility was measured by four indicants: frequency of change in the family's residence, expected period of occupancy for the new residence, duration of the husband's employment with one firm, and spatial distance involved in the most recent move; environmental transience was indicated by the frequency of pleasure travel trips. Social integration was not a designated focus of the original data collection efforts. It was, however, of major theoretic concern as a potential indicator of mobility-related dysfunctions. As a result, several

available measures, though imperfect, were employed as partial indicants of social involvement and integration. Integration into the larger (nonlocal) community was indexed by readership of nonlocal news-feature and/or sophisticated magazines (e.g., New Yorker, Punch). Involvement in the local community alone is measured by membership in various types of local social and civic organizations.

## Results

The findings presented in Table 1 concerning the frequency of exposure indicate that newspaper ads and real estate firms were the most important sources of information in the home buying process. Walking or riding through prospective residential areas and consulting friends followed as information sources utilized by a majority of buyers, while all other sources were used far less frequently. Cross-cultural differences appeared in the greater dependence of Connecticut buyers upon friends and co-workers vs. the use of relatives as information sources in Northwest England. The English buyers were significantly ( $p < .01$ ) more likely to mention employees of lending institutions, attorneys, and "For Sale" signs as useful sources of information in their decision process.

TABLE 2

Correlation of Social Integration and Mobility Measures with Information Source Sequence Indexes (Spearman Rank Correlation Coefficients)

Social Integration and Mobility Measures	Information Source Sequence Index (Order Contacted)							
	PSX		MMX		FCX		SIX	
	HTC (N=90)	NWE (N=63)	HTC (N=119)	NWE (N=81)	HTC (N=124)	NWE (N=91)	HTC (N=80)	NWE (N=81)
<b>A. Mobility</b>								
YRSOC: Yrs husband expects to occupy house	.01	-.09	-.20 <sup>b</sup>	-.17 <sup>b</sup>	.00	-.004	-.11	-.06
NMOV7: Number of moves in past 7 years	-.13	-.06	.06	.09	.07	-.10	.02	-.10
MBLTY: Mobility index, scaled 1-3	-.07	.00	.19 <sup>b</sup>	.19 <sup>b</sup>	.04	-.02	.08	.01
YRSFH: Years husband employed by firm	.17 <sup>c</sup>	.17 <sup>c</sup>	-.12	-.07	-.06	.06	-.09	.01
LOCPR: Distance moved, scaled 1-5	-.31 <sup>a</sup>	-.23 <sup>a</sup>	.06	.16 <sup>c</sup>	-.11	-.12	-.00	.13
NPTRP: No. of pleasure trips 100 miles or more last yr.	.06	.10	.06	.13	.16 <sup>c</sup>	.06	.06	-.08
<b>B. Social Integration</b>								
MED2: No. of new-feature mag. read reg. and/or subs.	-.02	-.12	.05	.03	.01	.16 <sup>c</sup>	-.06	.19 <sup>b</sup>
MED4: No. of sophisticated mag. read reg. and/or subs.	.07	.07	-.01	-.08	-.01	-.03	.01	-.03
MEDIA: No. of news-feature and soph. mag. read reg./subs. to	-.01	-.08	.04	-.01	.02	.11	-.06	.16 <sup>c</sup>
NNEWS: No. of national newspaper read reg. and/or subs.	.10	.04	.04	.26 <sup>a</sup>	-.11	.07	.04	.15 <sup>c</sup>
CVS01: Memberships of hus./wife in rel. or organization	.01	-.02	.03	.02	.03	-.08	.00	-.01
CVS02: Memberships in civic-PTA organization	.09	.11	.04	-.03	.00	.12	-.11	.09
CVS03: Memberships of fraternal societies	.01	.02	.20 <sup>b</sup>	-.04	-.06	.14 <sup>c</sup>	.14 <sup>c</sup>	.05
CVS04: Memberships in hobby, sports, garden clubs	.09	.02	-.13	.06	.02	-.04	-.00	.15 <sup>c</sup>
MBRSP: Total no. of above org. hus./wife belong to	.07	.04	.10	-.05	-.002	.09	.03	.10

a, b, c: Statistically significant beyond the .01, .05, and .10 level respectively.

The typical (median) sequence of contacting information sources tends to parallel the relative frequency of exposure. Newspaper ads were generally used as the first source in both countries, followed by real estate agents in Connecticut and relatives in England. Friends were usually contacted as the third source in both markets. In apparent contrast with the findings of many studies (e.g., Robertson, 1971), however, personal sources were typically used earlier in the decision process than most commercial sources or self-initiative sources. This may imply that home-buyers need to establish socially relevant criteria for evaluating housing alternatives (via personal sources) before they can effectively process the information from commercial sources. It is likely that personal sources are used throughout the buying process, but the measures discussed here consider only the first-use occasion.

As one might expect, the sources that were utilized more frequently also tended to be perceived as more helpful and thus worthy of recommendation to other buyers. The functions performed by commercial sources appear to be particularly appreciated, as the buyers recommended this channel of information with twice the frequency of any other channel. Cultural differences were most notable in the importance attributed to real estate firms and friends among Connecticut buyers, and in the commendation given to employees of lending institutions by English buyers.

In order to more precisely define the extent of relationships between the component indexes of physical mobility and community integration and the measures of information source utilization, Spearman rank correlations were computed. Portions of the resultant inter-correlation matrices are presented in Table 2 for one of the utilization variables: sequence in which the information channel was used. The variations in sample sizes indicated (from 80 to 119 for Hartford and from 63 to 91 for England) represent differences in sample attrition rates resulting from inapplicability of the sequence measure for respondents who did not use the source.

Examination of the statistically non-chance ( $p < .05$ ) correlations suggests several differences and similarities across the two social systems studied. For both the Connecticut and Northwest England samples, physical mobility of the environmental commitment type was weakly, but significantly, associated with later use of mass media sources. In contrast, increasing residential mobility, particularly as evidenced by distance moved, was generally associated with earlier use of personal sources. Additional analyses indicated that this, as anticipated, was most notable for one type of personal source, the co-worker. The importance of the job environments was further supported by the finding that occupational mobility was significantly related to the sequence in which personal sources were contacted.

The social integration measures yielded generally insignificant relationships for the Connecticut sample with the possible exception of some later use of media and self initiative sources by fraternal society members. No consistent pattern of source contact seemed to emerge, however. For the English sample, on the other hand, social integration, primarily as indexed by print media orientation but also in part by club membership, was related to some later use of media sources and of formal and self initiative sources. This may indicate some tendency for the highly socially involved to rely more on interpersonal contacts (rather than on their independent observations) early in the decision process as a means of identifying reliable commercial and media sources for later use. All correlations were fairly low, however, suggesting some lack of sensitivity

in the measures employed as well as the complexity of the relationships examined. Still, the finding that social integration and physical mobility might function differently as determinants (or correlates) of information search patterns is worthy of note. This might thus serve both as an indicant of possible cultural differences in source utilization (which should be taken into account by multinational corporate planners) and as an additional component of cultural life style variability (the determinants and consequences of which should be investigated by social scientists).

Although the corresponding data are not tabled here because of space limitations, certain similarities and differences across the two social systems were also noted for the other two measures of information source utilization. For example, residential mobility was positively related to use of ("exposure" to) self initiative sources in making housing decisions. It was also related positively to use of personal sources, but only for the Connecticut sample. Membership in civic groups was, as would be predicted, positively associated with self initiative source use and, somewhat more weakly, with personal source use. Fraternal or social club membership proved a significant correlate only for the English sample, however, having positive correlations with use of personal sources, formal commercial sources, and mass media sources. As indicated in Table 2, however, such overall use-related findings may occasionally mask sequence differences. For this same sample, fraternal organization membership was associated with later use of formal and self initiative sources and with earlier use of mass media sources.

As can be seen from the Table 2 data, correlational analysis indicated a relatively greater importance of the physical mobility variables as predictors of information utilization. In order to further examine the influence of these variables, the three indices of utilization (exposure, sequence and evaluation) were broken down by level of mobility. The results of these cross-cultural comparisons are presented in Table 3. The mobility segments represented in Table 3 were defined by average scores on three measures of environmental commitment (residential) mobility: number of moves in the past seven years, duration of occupancy of the last two residences, and expected duration of occupancy of the new residence. The measures were separately scaled and then summed to construct the mobility index on which the buyer samples were then trichotomized.

As can be seen from the tabled data, exposure to each of the information channel types was quite high in all three mobility segments. There was heavier reliance overall on mass media sources (a function perhaps of their availability) and certain commercial sources (a function of their close--and, for all practical purposes, required--association to the housing decision). There was also generally earlier use of the mass media sources.

Some cross-cultural differences may also be noted here. While the mobility pattern holds fairly constant (with more mobile buyers evidencing greater use of self initiative sources), the differences in absolute exposure levels suggest a greater utilization of self initiative channels by the English movers. In contrast, Connecticut movers showed greater use of personal and mass media sources and, for the low mobile group, for commercial sources as well.

Table 3 data also indicate the differences among the component indices of information source utilization.



TABLE 3

## Comparison of Information Utilization Indexes by Level of Mobility

Information Source Utilization Index	Level of Mobility					
	High		Moderate		Low	
	HTC (N=67)	NWE (N=52)	HTC (N=77)	NWE (N=115)	HTC (N=40)	NWE (N=43)
<b>A. Percent Exposed</b>						
Personal (EPSX)	76%	63%	68%	57%	70%	67%
Mass Media (EMMX)	90	85	87	77	90	79
Commercial (EFCX)	94	92	94	87	98	88
Self-Initiative (ESIX)	66	81	61	73	58	63
<b>B. Median Sequence of Exposure<sup>a</sup></b>						
Personal (PSX)	3.0	2.5	3.0	3.0	3.0	2.5
Mass Media (MMX)	2.0	1.0	1.0	1.0	1.0	1.0
Commercial (FCX)	3.5	3.5	3.0	3.0	3.5	4.0
Self-Initiative (SIX)	4.0	3.5	4.0	3.0	4.5	3.5
<b>C. Percent Recommending</b>						
Personal (RPSX)	37%	11%	23%	16%	20%	16%
Mass Media (RMMX)	33	21	31	26	45	37
Commercial (RFCX)	63	75	77	64	80	63
Self-Initiative (RSIX)	18	17	14	21	18	12

<sup>a</sup>The median order of contact (sequence) was calculated on the bases of those who were exposed to the source. The actual number of observations for these estimates can be determined by multiplying the subsample size (N) by the percentage shown for the exposure measure.

While the English sample showed higher usage and earlier usage of self initiative sources, they did not recommend them more highly. Both samples recommended commercial sources most highly, although they were typically employed in the decision process after mass media and personal sources (which may thus serve, in part, as the major identifiers of reliable commercial sources). As indicated previously, this latter function may be especially evident for the strongly socially integrated.

One interesting result that may reflect some cross-cultural difference in the impact of mobility was that commercial sources were more highly recommended by the more mobile buyers in England. This was not the case in Connecticut, where the less mobile buyers were most likely to recommend these commercial information sources. Further analysis of this pattern indicates that the high mobile English buyers tended to recommend attorneys, whereas the low mobile Connecticut buyers recommended builders. In England, the higher social status of the high mobiles may predict greater reliance on attorneys as a community contact. In Connecticut, the more stable low mobiles were perhaps more likely to be interested in having new houses built, and hence turned to builders as a major information source.

#### Implications and Conclusions

This research further demonstrates the rather apparent fact that both similarities and differences exist across identifiable social systems (in this case, across national boundaries). Of greater interest is the demonstration that marked variability exists across commonly

employed indexes of mobility and social integration. Of still greater heuristic importance is the notion that information source utilization is clearly multi-dimensional--that widespread use of a given source, for example, does not necessarily indicate early use of that source. This noted complexity of information-related behaviors has theoretic significance for researchers investigating information search patterns. It also has pragmatic implications for those involved in targeting information to consumers. Analysis of the relative importance of information sources measured in terms of exposure must be supplemented by consideration of the sequence in which sources are utilized. The design of efficient and effective communication packages (e.g., consumer information pamphlets) requires knowledge of how and where information modules might best fit into the ongoing information stream. Proper sequencing of information flows can facilitate the consumer's ability to cope with and even benefit from the "information explosion." Hopefully, the more actionable bases for segmentation, such as those discussed above, may provide certain of the visible components required to allow planners to utilize this complexity in reaching appropriate consumers, through appropriate channels, at appropriate times.

Since increasing rates of mobility usually accompany the industrial development of a society, the impact of residential change upon the family and the community should be considered in the broader cost-benefit analysis of growth policy. Knowledge of the role performed by various information sources in facilitating family adaptation to residential change is vital to the planning of community development programs. If mobility and social integration are considered to be important determinants of information seeking behavior in

residential choice, then the effects of the resultant source usage patterns upon other aspects of the decision process (e.g., choice criteria and post-purchase satisfaction) need further study before the functional or dysfunctional consequences of these two input variables can be fully evaluated. The influence of intervening variables, such as the motivation for moving and whether the change is perceived as voluntary or involuntary, may mediate the dysfunctional aspects of mobility, particularly of the environmental commitment type. For example, intercity moves that are associated with major job promotions may result in more limited information seeking and less scrutiny of alternatives because the family's economic outlook is affected by a halo of optimism. The perceived complexity of the residential change, affected by stage in family life cycle and accumulated moving experiences, is also likely to moderate the effects of mobility on information seeking behavior. In general, the effects of mobility and social integration upon information seeking in residential choice decisions should be examined as an environmental adjustment system which is closely tied to family satisfaction and the perceived quality of life.

#### References

Alan R. Andreason and Peter G. Durkson, "Market Learning of New Residents," Journal of Marketing Research, 4(May, 1968), 166-176.

Donald F. Cox, "Clues for Advertising Strategists," Harvard Business Review, 39(September-October, 1961), 160-176, and (November-December, 1961), 160-182.

Brenda Dervin and Bradley S. Greenberg, "The Communication Environment of the Urban Poor," in F. Kline and P. Tichenor (eds.) Current Perspectives in Mass Communication Research (Beverly Hills, California: Sage Publications, 1972), 195-233.

Herbert J. Gans, "Vance Packard Misperceives the Way Most Americans Move," Psychology Today, 6(September, 1972), 20-28.

Donald J. Hempel, 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).

Donald J. Hempel, "Search Behavior and Information Utilization in the Home Buying Process," 1969 Fall Conference Proceedings (American Marketing Association, 1969), 241-249.

Jack McLeod, Scott Ward and Karen Tancill, "Alienation and Uses of the Mass Media," Public Opinion Quarterly, 29(Winter 1965-1966), 583-594.

Vance Packard, A Nation of Strangers (New York: David McKay Publishers, 1972).

Thomas S. Robertson, Innovative Behavior and Communication (New York: Holt, Rinehart, and Winston, Inc., 1971).

Everett M. Rogers, Modernization Among Peasants: The Impact of Communication (New York: Holt, Rinehart, and Winston, 1969).

Everett M. Rogers and F. Floyd Shoemaker, Communication of Innovations (New York: The Free Press, 1971).

Jagdish N. Sheth and S. Parakesh Sethi, "An Approach to Developing a Theory of Crosscultural Buyer Behavior," (Paper presented at the American Psychological Association Conference, Montreal, 1973).

Alvin Toffler, Future Shock (New York: Random House, 1970).

Victor H. Vroom, "Industrial Social Psychology," in Gardner Lindzey and Elliot Aronson (eds.), The Handbook of Social Psychology (second edition), (Reading, Mass: Addison-Welsey), 5, 196-268.

## HIERARCHIES IN GOODS-CHARACTERISTICS ANALYSIS

Kelvin Lancaster, Columbia University\*

### Abstract

Two major problems in demand theory are those of justifying the analysis of demand for goods within a group without reference to goods outside the group, and constructing a model of consumer behavior that conforms to recognized limits on human information-processing abilities. This paper investigates the use of hierarchical structures in solving these problems.

### Introduction

The problem with which this paper is concerned is that of narrowing down the set of all characteristics possessed by goods within a group to a smaller subset of relevant characteristics, in terms of which the behavior of consumers with respect to choices over the group can be fully explained. The notion of relevance in this sense was introduced in Lancaster (1971) and is related to, but not identical with, the notion of "salience" in Fishbein (1971). Relevance is of particular importance in the economist's version of the characteristics model, in which the characteristics are, so to speak, looked at from the goods end rather than the consumer's end and thus consist potentially of all the physical, chemical, biological and other objective properties of the goods. This view was stressed in the original model of Lancaster (1966) and again, more recently, in Ratchford (1975). Behavioral models, looking at characteristics from the consumer end, have generally pre-selected the apparently relevant characteristics, but may ignore taken-for-granted properties.

It is worth emphasizing at this point that the economist has always taken a holistic view of the consumer as a decision-maker with a unified overall objective (maximization of utility, or attainment of a preferred state), so that all his decisions should be considered to be interconnected. Both this view and the most common behavioral view of the consumer as fragmented into a complex of more or less independent decision-making particles are obvious over-simplifications, but they are oversimplified from quite different directions. Thus the economist does not take it for granted that a particular group of goods can be separated off and consumer behavior with respect to this group be analyzed without reference to events outside the group, but considers it necessary to specify what special structural properties enable this separation to be made.

Goods may be separated into distinct groups on the basis of horizontal properties. Goods in two groups may, for example, possess the same essential characteristics but may be clustered so that the relative contents of different characteristics are comparatively close within each group, but widely different between groups (sports cars and stationwagons, for example). Our main interest here, however, is in vertical separation. This does not mean that the groups themselves are conceived of as standing in a vertical or hierarchical relationship, but that the separation between the groups can be established prior to any detailed investigation of the properties of the group.

### The Utility Tree

Leontief (1947) introduced the concept of a separable function, where  $F(x_1, \dots, x_n)$  is separable into  $s$  groups if the  $n$  variables can be partitioned into  $s$  groups and functions  $f^r(x^r)$  exist such that  $F$  can be written in the form  $[f^1(x^1), \dots, f^r(x^r), \dots, f^s(x^s)]$ , where  $x^r$  is the collection of variables assigned to the  $r$ 'th group. Each  $f^r$  is called a "branch" of the function  $F$ .

The importance of separability for a utility function, should it have that property, is that maximization of utility can be regarded as a consistent two-stage process of the kind first described by Strotz (1957), in which the consumer first makes an optimal allocation of his budget among the groups, then spends that allocation optimally within the group. After some debate in the literature, it has been established that the weak form of separability given above is sufficient for consistency of the two-stage process, due to the work of Blackorby et. al. (1970). If the utility function is separable, it is possible that one or more of the branch utility functions  $f^r$  may themselves be separable into sub-branch utility functions, and that some of these may in turn be further separable (giving us a real "tree"), implying the existence of a consistent multi-stage optimization.

If utility maximization can be regarded as a multi-stage process, the final stage will consist of choice among a restricted group of goods, subject only to the group budget. Effects arising from, say, price changes within another group will be very indirect and will be apparent only insofar as they affect the budget allocation of the group. Thus separability of the utility function would go a long way towards enabling us to analyse the demand for a group of goods in comparative isolation from events outside the group.

The traditional discussions of separability have been in terms of a utility function on goods, with goods as the arguments of both the overall utility function and the branch functions. In the characteristics approach, the utility function is conceived of as a function of characteristics quantities. Now there is no reason why the utility function on characteristics should not be separable, the groups now being groups of characteristics rather than of goods, but the multi-stage optimization property (which is what we are really interested in) no longer follows from separable utility alone.

The problem is that, although the utility function is defined on characteristics, the budget constraint is in terms of goods and characteristics cannot be purchased separately but only by buying the goods which possess them. Characteristics which may be separable in the utility function may not be separable in the budget because they are possessed by the same good. Suppose, for example, that there are three characteristics and the utility function is separable into three branches, one for each characteristic. If all goods possess all three characteristics, no division of the optimization process is possible and the separability of the utility function has no operational significance.

It is obviously sufficient for multi-stage optimization in a characteristics model that (a) the utility

\* The author wishes to acknowledge the assistance of the National Science Foundation (SOC-7514252).

function on characteristics be separable into groups of characteristics and (b) that the technological relationship between goods and characteristics be such that for every group of characteristics there be a group of goods which possess no characteristics other than those of the characteristics group. This requires not only that the consumption technology can be partitioned into groups of goods having characteristics exclusive to that group, but that this partitioning corresponds to the groups of characteristics into which the utility function is separable.

This double requirement on the partitions may seem to be a very strong one, but it is, of course, satisfied implicitly if the utility function on goods can in fact be shown to be separable, since the structure of the utility function on goods is a compound of the structure of the utility function on characteristics and the structure of the goods-characteristics technology. Stating the properties in terms of characteristics merely reveals how strong a property is that of separability in terms of goods, it does not of itself create the difficulties.

Separability is, of course, a question of fact. Unfortunately there is little in the corpus of empirical work on demand that provides any guidance as to whether individuals act as if their utility functions are separable or not. The two most used models in econometric demand studies, the "Rotterdam Model" and the "Linear Expenditure Model", both can be shown to implicitly assume strong separability of the utility function, along with a variety of other more or less heroic assumptions, and provide no test of any separability hypothesis. The most that can be said is that the assumption of separability (of goods into relatively broad groups) has not been shown to introduce any identifiable problems. For further details on the various econometric studies, the reader is referred to Brown and Deaton (1972) or Philips (1974).

In spite of the lack of any explicit confirmation from broad econometric studies, some kind of separability assumption is implicit in all attempts to study the demand for closely related goods, whether by economists or by others concerned with consumer research. It is simply taken for granted that we can sensibly study choices over a properly defined group of goods without reference to the details of choice over other groups.

#### The Group

Following through the discussion of separability in the characteristics model, we can see that the separation of certain goods into a group that can be studied in isolation requires an appropriate coincidence between the way in which the utility function can be separated and the way in which the technology of goods-characteristics relationships can be partitioned. If the partition boundaries coincide exactly, there is no ambiguity about the division into groups. If the partitions coincide only in some cases, then the groups are defined by a sort of highest common factor logic. That is, if  $C_1$  and  $C_2$  are groups of characteristics which are separable in the utility function, but characteristics from both groups are possessed by all goods in some group  $G_1$ , then we obtain only a single separable group (provided no goods outside  $G_1$  possess these characteristics), not two. Conversely if goods groups  $G_1$  and  $G_2$  possess no common characteristics, but the characteristics from the groups are not separable in the utility function, then the true separable group is that of the goods in  $G_1$  and  $G_2$  taken together.

Consider, for example, two food characteristics, "flavor" (taken to be one-dimensional for the sake of illustration only) and "calorie content". It seems reasonable to suppose that flavor, a direct effect on the senses, and calorie content, related to health and personal appearance, impinge on different branches of separable utility function. Because of their technological jointness in food, however, the separability of the utility function is not the determining influence in the grouping. Or consider a weight-watcher dinner and an exercise-parlor session, two goods with no objective characteristics in common but which possess characteristics likely to appear in the same branch of the utility function, and thus not in totally separate groups.

In the worst-case situation, the utility function may be separable into many branches and the consumption technology be partitionable into many groups of goods with their own sets of characteristics, but there may be no coincidences between the divisions and no groups of goods that can be fully separated from each other.

Due to the role played by the structure of the consumption technology in determining groups, grouping may be changed as a result of technological innovation, either embodied in new goods or arising from newly discovered information about the properties of existing goods. Air travel, for example, has consolidated the formerly distinct groups of goods (using the term in the general sense) "visits to Europe" and "annual vacation activities" into a single group, by removing the disparity in time required for the two sets of activities. Discovery of the carcinogenic properties of cigarette smoking was equivalent to a change in technology (even though cigarettes had always been considered vaguely unhealthy) and certainly changed the structure of the group by making the tar-nicotine content a highly relevant characteristic, although it may not have caused a regrouping. Technological change need not consolidate groups (as in the air travel example) but may also divide them. Modern audio technology permits separation of the activities "listening to a symphony" and "going to a concert", which were formerly always in combination.

#### Consumer Diversity

The preceding discussion has been confined to the analysis of how separability of an individual utility function, combined with an appropriate structure of the technology, enables that individual to optimize in more than one stage and thus enables him to consider groups in relative isolation. Obviously, the concept of the group is of no real use in the analysis of market behavior unless the separation into groups has essentially the same pattern over many, if not all, individuals.

The two-part structure of the characteristics model provides a much firmer basis for the existence of groupings which are common to many consumers than does the traditional economic model which sets up utility as a direct and individual function of goods. In the characteristics model, groups can only be separated along partition lines in the technology, and only then of course, if utility is separable in a corresponding way. Since the technology is common to all consumers, the grouping for any consumer is built out of the same basic blocks as those for any other consumer. This does not guarantee the existence of any groups which are common to all consumers since the technology may, for example, be partitioned into three distinct blocks of goods (no two blocks possessing the same characteristics) while some consumers have utility functions separable in such a

way as to make the first two blocks into one group, the third block into another, and other consumers group the second and third blocks together, with the first separate. In this case, all consumers see a pattern of two groups, but the groups are not the same for all. In spite of examples of this kind, it is apparent that the existence of common building blocks containing many goods makes the existence of common groupings a reasonable hypothesis, and makes it a certainty if we can assume that some of the separability properties of utility are common to all persons.

Virtually all econometric demand studies (perhaps all) that are based on some underlying model of consumer behavior and are not a set of ad hoc equations, implicitly assume that the aggregate demand can be treated as if it resulted from the decisions of a single representative consumer. Such models have no place for consumer diversity, which would nullify some of the properties assigned to the aggregate demand functions. Taking the goods groupings to be common to all consumers is then a straightforward consequence of non-diversity, providing no test of commonality.

We cannot, of course, analyse the fine structure of demand within a group without supposing that consumers are diverse. The hypothetical representative consumer in the demand studies is someone who brushes his teeth with every brand of toothpaste, in the proportions in which they are sold on the market, so that this approach has nothing to say about the choice among kinds of toothpaste. Insofar as economist's traditional demand studies are useful at all, it is in the light they can throw on very broad properties of demand - in which "goods" are aggregates, even complete groups, rather than individual products. Although these studies are commonly considered "micro" rather than "macro" economics, they are really a kind of half-way "macro-micro".

"Micro-micro" demand theory, that can handle product differentiates and not just aggregates like "automobiles", depends on taking account of the real diversity that exists between individuals. The more micro the analysis, however, the greater the dependence on the ability to separate individual groups of goods. If this cannot be done, we are forced to contemplate a model in which a consumer is making a single-stage choice over tens of thousands of different products.

If we put all this together, we see that a feasible analysis of demand at the most micro level really requires

(1) that all (or most) individual consumers share a common structure of preferences with utility functions that are separable in a universal way, so that there exist goods groups that are the same for different persons

(2) but that the individual preferences are diverse within this common structure, so that different individuals make different choices within the groups and allocate different proportions of their total budget to the various groups.

Although the assumption of diversity constrained by a universal structure is presupposing a great deal about the human psyche, that is what is implicit in all fine-structure analysis of demand over relatively small groups of products. Since analysis at this level has been carried out by those in market research for many years, we can presume that the underlying assumption is not unacceptable.

## Intragroup Analysis

We have now established the conditions which enable us to single out a group of goods and analyse the demand for goods within the group without explicit reference to goods in other groups, confident that this corresponds to the decision structure common to all individuals in the market yet leaves scope for diversity of choice within the group. Analysis of demand over a relatively small group of goods is clearly a much simpler problem than that of analysis over the set of all possible goods, but remains complex. We remain strongly interested in additional structural or other properties than permit further reduction of the problem.

If we take any set of goods likely to emerge as a group for decision-making purposes, and take the characteristics to be the set of all objective properties possessed by goods within the group, it is obvious enough that potential number of characteristics can be enormous. The number will vary a great deal - if the group is "automobiles" (or any subset of automobiles), the number of discernible properties is very large, while if the group is "soft drinks", the number will be much smaller. Nevertheless, analysis in terms of characteristics is potentially useful only if the number of characteristics actually entering the decision process can be greatly restricted. We shall refer to these as relevant characteristics.

Lancaster (1971) explored some of the grounds on which characteristics might be rejected as irrelevant on a priori criteria. These included universal characteristics (possessed by all goods, and possessed by goods within the group to approximately the same extent per dollar), invariant characteristics, characteristics made irrelevant through certain types of satiation, redundant characteristics (in fixed proportion to other characteristics). In most cases, however, we can expect that there remain a very large number of potentially relevant characteristics even after the maximum possible use of the preceding criteria.

Now consider the group from the other end, from the viewpoint of the consumer rather than the technology. We all know that we, as human decision-makers, are very limited as to the number of different variables that we can handle at one time. The work of Miller (1956) and others has validated this introspection experimentally, and come up with five to nine as the maximum number of variables. Fishbein (1971) takes the number of salient outcomes considered by a consumer to lie in this same range, presumably a number not chosen at random. Let us accept the magic number seven as expressing the limits of human decision-making powers, then we must conclude either that goods do not ever possess more than seven potentially relevant characteristics or that decision-making, even over the single group, occurs in such a way that not more than seven characteristics are being considered at any one time.

There are, of course, many studies which have investigated the number of relevant characteristics in particular markets at particular times, using a variety of methods to achieve results, including asking the consumers themselves in one way or another, using perceptual maps or principal component analysis, or analysing brand-switching data. Our interest here is not in the actual characteristics chosen as relevant in some context, but in the process by which the decision of relevance or nonrelevance is reached.

The simplest hypothesis is that of random selection - consumer simply becomes aware of certain characteristics

in a sequential and random way and closes his mind when he reaches his information-processing limit. This does not require diversity of preferences in order to generate different choices and would imply that the role of advertising was to make sure that the characteristic thought to be a good selling point was made prominent so as to get in before the door closed. Acceptance of such a model might be no big thing in marketing research, but would have profound consequences for economics because the foundation of all microeconomic policy is that consumer choice implies attainment of a preferred or constrained optimum position. The random model could lead to the conclusion that an expert might make a better choice for the individual than he could himself. This is certainly conceivable, but if true the market economy, even in the ideal perfect state so often modeled by economists, would lose its ultimate best claim to superiority over planning.

As an economist, the author is therefore biased in favor of some model that involves purposive choice at all stages, and thus towards the type of model in which there is a deliberate sequential choice process, each step of the sequence involving a consideration of an appropriately small number of characteristics.

#### Hierarchical Structures

A decision process in which a choice involving a restricted number of parameters is made, after which a further choice is made from another restricted set of parameters, and so on down the sequence, is necessarily hierarchical unless it is purely random. The ordering of the hierarchy determines which set of parameters is considered first, second, and on through the sequence.

More than a century ago, Karl Menger discussed a model of the consumer in terms of a hierarchy of wants. He was concerned with a sequence of wants of diminishing importance, each satisfied by successive allocations of the same good rather than by different goods, and used the model primarily to establish the notion of marginal utility. This is not the kind of hierarchy which we need in the present context, but comes much closer to it if the successive "wants" are thought of as being satisfied by different characteristics of the same good.

A more recent hierarchical model, and one that is directly comparable with the characteristic model, is that given by Ironmonger (1972). It should be noted that the Ironmonger work, although published much later than the original Lancaster (1966) paper, was developed independently and much had been embodied in an unpublished 1961 dissertation. The Ironmonger analysis is cast in terms of wants rather than characteristics, but the effective variable in his utility function is the "number of units of satisfaction of want *i*" which is simply the view from the consumer's end of the number of units of characteristic *i* obtained from consumption. Ironmonger's utility function is a strictly hierarchical one, with wants completely ordered (a lexicographic preference order with satiation) so that the consumer's decision at any one stage is confined to that of satiating the next unsatisfied want in the ordering. Interesting results follow from the analysis solely because goods possess multiple characteristics and thus concentration on a single want at a particular stage results in other wants being partly satisfied. The multiple characteristics properties of goods generate many results similar to those of Lancaster (1966, 1971), in spite of the great differences in the assumed structure of the preferences. (The standard Lancaster analysis assumes characteristics act simultaneously, not sequentially, on utility).

Two aspects of the Ironmonger structure seem to rule it out as far too extreme for our present purposes. A fully hierarchical preference ordering would imply that a consumer would never think about clothes until after he had eaten (assuming the ordering was in that direction), while separable but nonhierarchical preferences would merely imply that the consumer could think about clothing without thinking about eating and vice versa. The latter seems surely to conform more accurately to reality. The other doubtful property is that of having only one characteristic at a time enter in to the decision process.

The utility model that fits best the reality of limited information processing ability and the various criteria for micro analysis of demand, together with the property that choice reveals preference, is one with a mixed structure. It should be separable, with the same general lines of separation for all individuals, but without hierarchical relations between the groups as a general pattern. Within the groups, there should be some hierarchical structure so that the final decisions can be made in several stages, in each of which only a few (but typically more than one) characteristics are taken into account.

Consider a potential scenario for choosing an automobile, assuming that the separability properties have already enabled this choice to be made without reference to goods other than automobiles. Stage I might consist of assessing the characteristics associated with whether the car actually runs or not - does it have an engine? Wheels? does the engine turn over? This is a simple blocking stage, and cars which are not accepted here will not be chosen whatever their other characteristics. (We assume repair is not possible). Stage II might consist of considering those characteristics that make the car suitable for family use - how many will it seat? how much baggage will it hold? Finally, Stage III might consist of comparing details in aesthetics, convenience, comfort, mechanical operation and so on.

This scenario illustrates the main features that would have to be taken into account in devising a multi-stage decision model, which we can summarize as follows: (1) How many stages are there, and is some decision about the number of stages itself a first stage of the process? (2) What determines the ordering of the stages, and is the determination of this ordering part of the decision process? (3) Finally, what kind of decision is made at each stage? Simple accept/not accept, for example, or some assignment of an actual utility number?

Once again, the special interest of the economist in the consumer as a whole needs to be stressed. An ad hoc model of a limited decision process that implied something about the consumer's overall preference structure or utility function that was very different from that usually assumed would find little acceptance among economists, unless it could be shown compellingly that the truth and generality of the model justified scrapping accepted economic theorems.

We should distinguish between the technique and the substance of the decision-making, since it is only the latter that affects the fundamental structure of the implied utility function. The number of stages, for example, is not important, but the kind of decision made at each stage can have a fundamental bearing on the implied preference ordering.

Consider the pass/fail kind of decision implied in Stage I of the suggested scenario given above, in which a subset of characteristics is checked against a minimum acceptable level and the good is not considered

further unless it passes on all counts. If this is followed by an optimizing choice over those goods which have passed, the implied preference structure is that of an incomplete ordering which is, however, well-behaved over those collections actually ordered. The ordering is incomplete because goods which fail the screening stage are not further considered and thus not ranked with respect to each other. But the final choice can be considered to represent a properly preferred position because it is the preferred choice from the final set, and all goods (or goods collections) in the final set can be taken to be preferred to all goods which failed the screening. Thus a simple pass/fail stage in the decision process does not cause any fundamental upset in the basis of preference.

Stage II in the example given might be taken as a more advanced type of pass/fail decision. Instead of a series of characteristics which all must be satisfied in order to pass the screening, the various characteristics are assumed to contribute jointly to some index of suitability as a family transport, as characteristics contribute jointly to utility in a utility function. A car passes this stage if the index has at least some minimum value. Again, if this stage is followed by a pure maximization stage, the final choice is a true preferred position. A screening stage of this kind does, however, introduce problems of both incompleteness and continuity into the preference structure.

Sequential sorting processes have been suggested as appropriate models for market research purposes - the Hendry model, for example (see Herniter (1974)), and so have various models based on threshold effects. For market research it is, of course, of the highest importance to determine which particular characteristics are taken into account at which stage. For our more general approach, it is important only that an individual be consistent in the structure of his decisions. If the group is well-defined, the number and specification of decision stages can be taken as individual, contributing to the assumed diversity of preferences over the group. We do not need a general and universal theory of the order of sorting.

In the pure screening model, only the last decision stage is a true maximizing stage. In this final stage, presumably, the consumer chooses his preferred collection of those characteristics left to this stage from among those goods whose other characteristics have been sufficient to pass the previous screening. The process is that of maximizing a (branch) utility function of certain characteristics, subject to obtaining minimum levels of other characteristics or functions of these.

Although the screening model has many attractions, has been used as the basis for market behavior models, and can be made to fit in reasonably well with the most important assumptions on preferences usually made by economists, it is not perfect. One is uncomfortable, in particular, with the status of above-minimum characteristics levels at the screening stages, which fail to count in the final decision. If a car is especially suitable for family use, why should not that factor count at the final stage, even to the extent of resulting in the choice of that car over another that actually has more final stage characteristics?

One potential modification of the pure screening model is to introduce some concept of "carry-over", where properties above the minima and one stage are carried forward to be weighed together with the next stage characteristics. If above-threshold quantities of individual characteristics are carried forward, we lose the property of limiting the number of variables per stage, which was the object of introducing multi-stage analysis.

However, we could carry forward the above-threshold level of some index (like family suitability in the example given) and still achieve reduction of variables.

Why not carry forward the whole value of the early stage indices, so that the final decision was a maximization of a function of indices, subject to minimum levels of specific indices? This could be done on the basis of a kind of separability of the utility function alone, and would preserve most of the properties liked by economists. But, although it reduces the number of characteristics considered at any one time (since the making up of each index is a separate operation), it does not reduce the number of goods - all goods in the group would be considered together in the final judgment. Thus it seems desirable to retain the essentials of goods-screening process.

Such a model, if it could be developed into an applicable form, would seem to come close to satisfying both the economists and the behaviorists. Until this, or some other model with equivalent properties, is fully worked out, economists and those concerned with detailed market behavior are likely to have somewhat different views on the use of characteristics analysis. Economists, interested in the broader analysis of market structures - as typified, for example, in Lancaster (1975) - will continue to look at characteristics from the objective, goods point of view. Market behaviorists will continue to look at characteristics from the subjective, consumer's end, as in Howard and Sheth (1969), Wilkie and Pessemier (1973), Fishbein (1971).

#### References

- Blackorby, C, Lady, G, Nissen, D and Russell, R (1970), "Homothetic separability and consumer budgeting", Econometrica, 38 (1970), 468-472.
- Brown, J and Deaton, A, (1972), "Surveys in applied economics: models of consumer behaviour", Economic Journal, 82 (1972), 1145-1236.
- Fishbein, M (1971), "Some comments on the use of 'models' in advertising research". In European Society of Market Research, Proceedings: Seminar on Translating Advertising Theories into Research Reality, 1971, 297-318.
- Herniter, J (1974), "A comparison of the entropy model and the Hendry model", Journal of Marketing Research, 11 (1974), 21-29.
- Howard, J. and Sheth, J. (1969), The Theory of Buyer Behavior, New York, Wiley, 1969.
- Ironmonger, D. (1972), New Commodities and Consumer Behaviour, Cambridge, Cambridge University Press, 1972.
- Lancaster, K. (1966), "A new approach to consumer theory", Journal of Political Economy, 74 (1966), 132-157.
- \_\_\_\_\_ (1971), Consumer Demand: A New Approach. New York, Columbia University Press, 1971.
- \_\_\_\_\_ (1975), "Socially optimal product differentiation", American Economic Review, 65 (1975), 567-585.
- Leontief, W. (1947), "Introduction to a theory of the internal structure of functional relationships", Econometrica, 15 (1947), 361-373.
- Miller, G. (1956), "The magical number seven, plus or minus two: some limits on our capacity for processing information", Psychological Review, 63 (1956), 81-97.
- Philips, L. (1974), Applied Consumption Analysis. Netherlands, North-Holland, 1974.
- Ratchford, B. (1975), "The new economic theory of consumer behavior: an interpretive essay", Journal of Consumer Research 2 (1975), 65-75.
- Wilkie, W. and Pessemier, E. (1973), "Issues in marketing's use of multi-attribute attitude models", Journal of Marketing Research, 10 (1973), 428-441.

BEYOND CONJOINT MEASUREMENT:  
A METHOD OF PAIRWISE TRADE-OFF ANALYSIS

Richard M. Johnson, Market Facts, Inc.

Abstract

This paper discusses data gathering and estimation procedures which overcome some of the problems encountered in conjoint measurement applications, and which lie on the boundary between conjoint measurement and more traditional techniques.

Background

Conjoint measurement has had a substantial impact in the relatively brief period since its introduction to the consumer research community (Green and Rao, 1971). One reason for this appears to be its capability of producing relatively "sophisticated" results, typically scaled at the interval level, from rather "primitive" data, normally consisting merely of rank order or paired comparison preference data. This strength allows the researcher to make reasonable predictions of choice behavior, even at the level of the individual consumer.

As we shall see, it is not a simple matter to decide in some borderline cases whether a particular procedure belongs to the class of conjoint measurement procedures or not. However, conjoint measurement procedures, as used in the study of consumer preferences, appear to have certain elements in common.

The first common element is that objects are conceptualized as consisting of "bundles of attributes" or, more precisely, bundles of attribute levels. As a method of collecting data we present several such bundles of attribute levels to respondents and gather information concerning overall preference among them.

The second common element is that of a "composition rule" according to which consumer preference is assumed to be affected by a number of unobservable variables. Frequently we assume that a consumer has a personal utility value associated with each level of each attribute, and that his degree of liking for a particular product is composed in some way from the utilities of its individual attribute levels. The composition rule most frequently assumed is the simple additive one, in which the overall utility of a product is assumed to be the sum of the utility values of its attribute levels.

The third common element is the estimating procedure by which an individual's utilities are inferred from preference data. Usually in consumer research the respondent is presented with a number of hypothetical product concepts and asked to rank them for preference. We ordinarily assume that his observable rank order of preference is monotonically related to the

sum of his unobservable overall utilities for the attribute levels represented in each hypothetical concept. The estimation problem is then simply that of finding estimates of individual attribute utilities such that when subsets of these are summed, these sums have the indicated rank order. Thus, specifically, the third element common to conjoint measurement applications has been estimation by nonmetric or "merely-order-preserving" algorithms, such as Monanova (Kruskal, 1965), monotone regression (Johnson, 1975), or linear programming (Srinivasan and Shocker, 1973).

We shall review data collection methods which have been used in consumer research applications, and propose a new method which overcomes certain practical problems with existing procedures. This new method leads naturally to other estimation techniques which will subsequently be described.

Methods of Data Collection

Within the field of consumer research, two rather dissimilar methods of collecting data have been used extensively, although several studies have used variations and mixtures of these. In order to distinguish among these as well as to describe the new method to be proposed, we shall consider an example.

Suppose we were studying consumer preference for table model radios, and we were to agree that such products could be characterized by six attributes, having the levels shown in Table 1.

TABLE 1  
One Consumer's Utilities

		<u>Utility</u>
1. <u>Brand</u> : . . . . .	G.E. . . . .	15
	Motorola . . . . .	13
	Sony . . . . .	3
	K-Mart . . . . .	0
2. <u>Type</u> : . . . . .	AM/FM/SW . . . . .	100
	AM/FM . . . . .	74
	AM only . . . . .	0
3. <u>Cabinet</u> : . . . . .	Wood . . . . .	30
	Wood-like plastic . . . . .	13
	Plastic . . . . .	0
4. <u>Sound Quality</u> : . . . . .	Superior . . . . .	38
	Adequate . . . . .	0
5. <u>Clock</u> : . . . . .	"Snooze" alarm . . . . .	78
	Regular . . . . .	33
	None . . . . .	0
6. <u>Price</u> : . . . . .	\$29 . . . . .	83
	\$49 . . . . .	39
	\$69 . . . . .	19
	\$89 . . . . .	0



The numerical values in Table 1 are utilities for a particular respondent. These are not usually known, of course, since our purpose is usually to estimate them from individual preference data. With an additive composition rule we are free to add or subtract a constant from the values for the various levels of any attribute. For simplicity these utilities have been arbitrarily scaled so that the least liked level for each attribute has a value of zero, and so that their maximum is 100.

One method of gathering preference data involves presenting the respondent with a number of imaginary product concepts, where each concept has a specified level for every attribute. Data collection methods of this type have been advocated by Green and Rao (1971), Green and Wind (1973), and Green and Devita (1975). The distinguishing aspect of this method is that it presents the respondent with entire concepts, specified with respect to all attributes. Although there seems to be no name consistently associated with this approach in the literature, we shall call it a "concept evaluation" method. For example, four imaginary radios that might be presented are shown in Table 2, together with a calculation of overall utility for our respondent.

The estimation problem is usually that of inferring a set of individual attribute utilities from observed preferences. In this example we have turned the process around to show how overall utility for each radio would be computed, given the individual attribute utilities in Table 1. Since Radio A has the highest utility it should be preferred to the others, which should in turn be ranked in order of their overall utilities.

TABLE 2

Overall Utility Calculations for 4 Hypothetical Radios

	<u>Radio A</u>	<u>Radio B</u>
<u>Brand:</u>	Motorola . . . (13)	G.E. . . . . (15)
<u>Type:</u>	AM/FM . . . (74)	AM/FM/SW . . . (100)
<u>Cabinet:</u>	Wood . . . . . (30)	Wood-like plastic. (13)
<u>Sound:</u>	Adequate . . . ( 0)	Superior . . . . . (38)
<u>Clock:</u>	Snooze alarm. (78)	None . . . . . ( 0)
<u>Price:</u>	\$49 . . . . . (39)	\$69 . . . . . (19)
<u>Overall Utility:</u>	. . . . (235)	. . . . . (185)
	<u>Radio C</u>	<u>Radio D</u>
<u>Brand:</u>	Sony . . . . . ( 3)	K-Mart . . . . . ( 0)
<u>Type:</u>	AM/FM . . . (74)	AM . . . . . ( 0)
<u>Cabinet:</u>	Wood . . . . . (30)	Plastic . . . . . ( 0)
<u>Sound:</u>	Superior . . . (38)	Adequate . . . . . ( 0)
<u>Clock:</u>	None . . . . . ( 0)	Regular . . . . . (33)
<u>Price:</u>	\$89 . . . . . ( 0)	\$29 . . . . . (83)
<u>Overall Utility:</u>	. . . . (145)	. . . . . (116)

It is not clear how many concepts must be presented to a respondent in order to permit robust estimation of his utilities, but it would seem prudent that the number

of concepts presented be on the order of three or four times the number of parameters being estimated. If there were six attributes with a total of 19 levels, as in the example, then the number of parameters being estimated is  $19 - 6 - 1 = 12$  (since we can arbitrarily set one value at zero for each attribute, and can scale all values to have arbitrary maximum). Therefore we might wish to expose the respondent to 40 or 50 concepts, each of which has a specified level on each of the six attributes. If we were dealing with 10 attributes, each with 4 levels, we might wish to present the respondent with ninety or more concepts, each specified with respect to ten attributes.

A second way of gathering data, called "Trade-Off Analysis" by Johnson (1972), has also been described by Fiedler (1972), Davidson (1973), Ross (1975), and Johnson (1975). The same technique appears to have been developed independently by Westwood, Lunn, and Beazley (1974).

This approach presents the respondent with attributes two at a time and asks for a rank order of preference for all the imaginary product concepts that could be described as combinations of their levels. The task is frequently made more systematic by presenting the attributes as rows and columns of a "Trade-Off Matrix," as shown in Table 3. The respondent would see Table 3 with no numbers, and fill in his rank orders in the cells. The cell entries are rank orders of preference that would be expected based on this respondent's individual attribute utilities for Type and Clock as shown in Table 1 and repeated in Table 3. The cell values in parentheses are sums of row and column utilities, and the preference ranks are consistent with these. The estimation problem ordinarily consists of finding a set of individual attribute utilities so that their pairwise sums, such as indicated in the cells of Table 3, have the desired rank orders.

In a study involving six attributes, as in the example, the respondent could be shown all fifteen of the possible matrices. With 10 attributes it would be more difficult to expose the respondent to all possible pairs, but one might pair each attribute with a subset of others so that about 20 matrices were filled out. If all attributes had three levels this would involve the presentation of  $9 \times 20 = 180$  stimuli to each respondent, about 9.5 times the number of parameters to be estimated. A procedure for choosing the specific pairs of attributes to be presented is suggested by Johnson and Van Dyk (1975).

TABLE 3  
A Trade-Off Matrix

Type	Clock		
	Snooze (78)	Regular (33)	None (0)
AM/FM/SW (100)	1(178)	3(133)	5(100)
AM/FM ( 74)	2(152)	4(107)	7( 74)
AM Only ( 0)	6( 78)	8( 33)	9( 0)

## Limitations and Shortcomings

A virtue of the concept evaluation method is that the respondent's task is relatively "realistic," since the stimuli to which he is reacting are complete concepts and he is not required to maintain an "all other things being equal" frame of mind. However, this realism is bought at the price of a considerable burden in interview length and task complexity when the number of attributes is greater than six or eight. Indeed, the magnitude of the respondent's task increases quadratically with the number of attributes, since larger numbers of attributes require not only that more concepts be presented, but also that each be more elaborately specified. By contrast, the trade-off matrix method has a clear practical advantage in studies with many attributes, simply requiring that the respondent fill out more matrices, a task that increases more nearly linearly as the number of attributes increases. Using the matrix format, studies involving a dozen attributes are feasible if the number of levels is around forty.

However, the matrix approach has a disadvantage which is avoided by the concept evaluation approach; it supplies a high degree of potentially artificial structure to the task. Since the stimuli are arranged neatly in rows and columns it is possible for the respondent to adopt a superficial "patterned" mode of responding by simply rank ordering the matrix cells from left to right or top to bottom. This behavior is observed particularly with respondents who profess a low degree of interest for the task or who find themselves overcome by its rather abstract nature.

A different but equally damaging type of lexicographic behavior is also observed with the concept evaluation approach. Here respondents who are disinterested or overloaded by the large amount of information provided for each concept sometimes narrow their focus, forming their preferences on the basis of presence or absence of one or two prominent levels.

It is clear that both tasks are comparatively difficult for respondents unused to thinking abstractly or processing large amounts of verbal information. We therefore have considerable motivation for finding data collection methods which provide the input for conjoint analysis at the level of the individual respondent while still being simple enough to be appropriate for the wide variety of respondents encountered in consumer research studies today.

### Pairwise Trade-Off Analysis

When using conjoint measurement with an additive composition rule, the role of the data may be regarded as that of providing inequalities which the calculated utilities should satisfy. For example, if a respondent prefers an AM/FM radio with no clock to an AM only radio with a clock, we know that the sum of his utility for AM/FM and his utility for no clock should be greater than the sum of his utility for AM only and his utility for a clock. Similarly, the statement of a preference for one concept over another, when they are both specified with respect to six attributes, merely provides one inequality, although it involves sums of

six values each.

It is clear that the minimal task that a respondent can perform while providing inequality information about his underlying utilities is to state which of two cells in a trade-off matrix he prefers. Of course, it is not necessary that he actually be shown a matrix. Individual pairwise questions may be asked in a format like the following.

Which would you prefer . . .

AM/FM radio with no clock	OR	AM radio with a clock
------------------------------	----	--------------------------

There is no reason why each stimulus element need consist of only two attribute levels, since this format could as well be used to elicit preference information for concepts specified on many attributes. To do so, however, would require more reading and thought on the part of the respondent and would still result in only one inequality per question.

The conceptual simplicity of the pairwise task renders it capable of being used productively with respondents from a wide range of socioeconomic levels, and our empirical research to date has shown that it produces data which are more consistent internally and of higher quality in general than either the concept evaluation or the matrix format approaches. However, this simplicity is obtained at the price of decreased efficiency in data collection. Pairwise trade-off interviews tend to be longer and somewhat more tedious than those using the matrix format approach.

A rank ordering of  $N$  stimuli by the concept evaluation method produces  $N(N-1)/2$  inequalities. Although only  $N-1$  of these are independent, and thus responses to  $N-1$  pairs could theoretically produce the same information under ideal circumstances, this will scarcely ever happen in practice. Such an ideal situation could occur only when a respondent's preferences are known before the pairs to be presented are selected. This might be approached in interactive data collection environments, where pairs presented late in the interview could be selected on the basis of preferences revealed earlier, but it is unlikely to occur when questionnaires must be printed in advance of the interview.

In comparison to the matrix format, it is conceivable that two pairwise questions could provide as much information as a ranking of all cells in a  $3 \times 3$  matrix. However, this will only occur if one attribute dominates another and we are fortunate enough to present the appropriate pairs. In general more information results from having a respondent rank matrix cells than can be obtained from the pairwise approach with reasonable interview length.

This means that the pairwise approach is not really feasible in studies with large numbers of attributes, where matrices would probably be preferred if the respondents are able to cope with the more demanding task.

## Special Considerations for Data Collection

The pairwise trade-off approach removes a number of restrictions which have previously applied to applications of conjoint measurement. The technique can be self-administered successfully, which makes it feasible for mail panel research, airline in-flight surveys, group administration, etc. If the researcher is willing to forego analysis at the individual level and can divide the stimuli among equivalent subsamples of respondents, the technique becomes feasible for telephone interviewing as well. In the limit, each respondent could be presented with one pair.

The questions of how many pairs to present and how to select these are complex, and the answers are not fully known. There are a few guidelines, however, which help in designing a questionnaire.

First, the number of pairs presented should probably at the very least be about three times the number of parameters to be estimated (number of levels minus number of attributes minus one). This is only the crudest of guidelines, however, since "easy" pairs contribute less information than harder pairs. The value of hard, as opposed to easy, pairs must, however, be considered in the light of respondent fatigue and attitude. It may be true that a sprinkling of very easy pairs will help keep the respondent motivated.

Second, it is essential that the design have a high degree of "connectedness" (Johnson and Van Dyk, 1975). While it is not necessary that every possible combination of attribute levels appear, it is certainly essential that every level appear in comparison with enough others to permit indirect comparisons of all differences not actually presented to the respondent.

The most interesting method of collecting data involves a computer-interactive environment where the computer administers the pairs, always presenting one which is nearly optimal in the sense of providing needed information given the pattern of responses obtained thus far. My colleague Frank Goode will discuss such an approach shortly.

## Computational Considerations

A surprising aspect of pairwise trade-off analysis is that, depending on how one views it, the estimation of utilities may be considered as an example of conjoint measurement, classical statistical inference, linear optimization, or stochastic modelling. Let us consider each viewpoint briefly.

In each case we shall conceptualize the computational problem as follows. Suppose that a respondent is given  $N$  stimulus pairs. Let there be  $n$  attributes, possessing a total of  $m$  levels. Consider a "design matrix"  $X$  of order  $2N$  by  $m$ , with elements of zero or one. The rows of  $X$  are considered in pairs, and labelled  $1l, 1r, 2l, 2r, \dots$  etc. The first row of each pair is associated with the lefthand element of that stimulus pair, and the second row of each pair is associated with the righthand element of that stimulus pair. A given row of  $X$  is entirely zero except for ones in the positions corresponding to attribute levels

possessed by that element of a stimulus pair. Consider also a vector  $Y$  of length  $2N$  with rows corresponding to those of  $X$  and which contains values of  $+1$  for the preferred element of each pair and  $-1$  for the nonpreferred element.

### 1) Estimation by Conjoint Measurement

One way to conceptualize the estimation task is as that of finding a weight for each column of  $X$  so that the weighted row sums, which may be regarded as a "prediction" of the vector  $Y$ , are as close as possible to  $Y$  in a particular sense. Consider the weights as elements of a vector  $W$  of length  $m$ . Then let  $XW = \hat{Y}$ . We simply wish  $(\hat{y}_{ir} - \hat{y}_{il})$  to have the same sign as  $(y_{ir} - y_{il})$ . This problem is none other than the familiar monotone regression problem, where order comparisons are restricted to subsets of size two. The computation may be handled by any monotone regression algorithm (Kruskal, 1965 or Johnson, 1975).

### 2) Classical Statistical Inference

We may regard the columns of  $X$  as "independent variables" and the vector  $Y$  as a "dependent" variable and use ordinary least squares procedures to find a weighting vector  $W$ . It is perhaps aesthetically preferable to regard this as a two-group discriminant analysis to which it is algebraically equivalent. In either case a set of column weights is sought, so as to maximize the difference between mean overall utilities for preferred and nonpreferred stimulus elements, subject to a constraint on the magnitudes of the weights themselves.

Alternatively, we might subtract adjacent rows of  $X$  from one another to get  $N$  rows corresponding to differences between preferred and nonpreferred stimulus elements. This would produce a matrix  $Z$  of differences of order  $N$  by  $m$ . Then we might seek a set of weights for the columns of  $Z$  which would produce row sums as "positive" as possible, in some sense. One criterion which might be utilized is that the sum of squares of the weighted row sums be maximized, subject to a constraint on the sum of squares of the weights. This is none other than the well-known "eigen problem," and the desired weights are simply the first eigenvector of  $Z'Z$ . (This criterion will be strictly appropriate only when weighted row sums are all positive, but may be close enough most of the time.)

### 3) Linear Optimization

Methods of linear programming can be used in the obvious way. Each row of  $Z$  represents a constraint in a linear programming tableau with  $m-n-1$  unknowns. We can add  $N$  "slack" variables to the system to insure nonnegativity of each row sum and seek to minimize the sum of these (Srinivasan and Shocker, 1973).

### 4) Stochastic Modelling

One way in which the pairwise trade-off method is unlike other conjoint measurement data collection methods is that it consists of a number of binary choice events which can, in a sense, be regarded as independent of one another. This suggests that it might be fruitful to devise some measure of fit of the model to the data which is more "probabilistic" than those

which have arisen in the nonmetric scaling literature. One obvious measure may be obtained by borrowing from the tradition of maximum likelihood estimation, with use of the logistic transformation.

As before, let the elements of  $XW = \hat{Y}$  be "overall utilities" for elements of the stimulus pairs. In particular,  $\hat{y}_{il}$  is the respondent's overall utility for the lefthand element of the first pair and  $\hat{y}_{ir}$  is his overall utility for the righthand element. We are interested in finding weights,  $W$  so that differences between adjacent elements of  $\hat{Y}$  have proper signs.

Let us define the vector  $U$  of length  $2N$  with elements

$$u_{il} = e^{\hat{y}_{il}} \quad \text{and} \quad u_{ir} = e^{\hat{y}_{ir}} \quad (1)$$

The elements of  $U$  are obtained by an exponential or antilogarithmic transformation of corresponding overall utilities,  $\hat{Y}$ . The values in  $U$  are all positive, and may be regarded as "multiplicative" analogs of the "additive" values in  $\hat{Y}$ .

Now, let us assume that the probability with which a respondent will choose the lefthand element of the  $i$ th pair is simply

$$P_{il} = \frac{u_{il}}{u_{il} + u_{ir}} \quad (2)$$

It is easy to show that

$$P_{il} = \frac{1}{1 + e^{(\hat{y}_{ir} - \hat{y}_{il})}} \quad (3)$$

commonly known as the logistic transformation.

Then a reasonable criterion of fit might involve a likelihood-type function of these estimated probability values. Specifically, let

$$\left. \begin{aligned} \delta_{il} &= 1 \text{ if the left member of the } i\text{th pair is preferred, } 0 \text{ otherwise.} \\ P_i &= \delta_{il} P_{il} + (1 - \delta_{il})(1 - P_{il}). \end{aligned} \right\} \quad (4)$$

Then  $p_i$  is the likelihood of the respondent's choosing as he did on the  $i$ th pair, given the probabilistic assumptions just stated.

The criterion we have adopted is the root likelihood, defined as

$$RLH = \left( \prod_{i=1}^N P_i \right)^{1/N} \quad (5)$$

This is the geometric mean of all  $N$  of the individual pairwise likelihood values.

The title of this paper contains the phrase "beyond conjoint measurement," which is motivated by the existence of this estimating method. As pointed out above, a common element of conjoint measurement procedures has been estimation by "merely-order-preserving" or nonmetric algorithms. These algor-

ithms produce solutions with as few "large" order violations as possible, and are satisfied by any solution which has no order violations. The likelihood criterion is quite different. It is not basically concerned with order violations but rather with accounting for data in a likelihood sense. There is no discontinuity in its "satisfaction" with a pair when the probability with which it "predicts" that response passes from below .5 to above .5.

This criterion has several desirable properties not shared by other indices of fit commonly used in conjoint measurement studies.

1) Like the others, it can be optimized by gradient-type computing methods. Unlike some others, however, its surface appears to be smooth and not to possess local optima. This removes one of the gravest problems in large-scale applications of nonmetric estimation methods.

2) In additive conjoint measurement the solution is determined only up to an arbitrary multiplicative constant. This source of arbitrariness is removed with the likelihood criterion since the scaling of a solution depends on its degree of fit to the data. If no well-fitting solution exists the estimates of utility for various attribute levels will be quite uniform. On the other hand, if there is very little error of fit the utilities will vary greatly from one another. Thus, an individual who gives coherent, reproducible data will have estimated utilities which are scaled quite differently from those of an individual who responds less consistently.

3) Merely-order-preserving methods are satisfied with solutions which reproduce order perfectly or nearly so. Kruskal's stress and Johnson's theta both tend to "stop worrying about" an order relation when violations become small. Accordingly, these methods tend to produce solutions with many ties or near-ties in differences among estimated utilities. The likelihood criterion, on the other hand, "continues to care about" each order relation, even after it is satisfied. Whereas a contribution to the gradient is proportional to the size of an order violation in those other methods, with the likelihood criterion a pair's contribution to the gradient is proportional to  $1 - p_i$ . Thus, although not actually a conjoint measurement estimating method, the likelihood criterion seems to provide better fits in terms of the number of violations than the conjoint measurement methods themselves.

4) One computational output is an estimate of the likelihood that the respondent would choose either member of any possible pair. These values have considerable potential usefulness in interactive data collection. Unasked pairs with estimated  $p_i$  values near either zero or one would probably not be worth asking. On the other hand, unasked pairs with  $p_i$  values near .5 are those about which information is lacking, and these are the pairs for which additional information would be of greatest value. Similarly, pairs already presented with low  $p_i$  values may have been answered in error and could perhaps be reasked.

## Summary

Pairwise trade-off-analysis has two desirable properties. First, the task presented to the respondent is the simplest one possible: a sequence of binary choices. Common sense and experience to date both suggest that a far wider range of people should be able to perform it successfully than is true of either the concept evaluation or the matrix format approach.

Second, data of the type gathered by this technique should be analyzable by a wide variety of computational approaches, including classical least squares techniques. Experience to date suggests that the maximum likelihood approach in particular has a number of desirable properties. These are

- 1) Invulnerability to local optima
- 2) A natural and non-arbitrary scaling for the solution
- 3) Tendency to avoid ties among sums of subsets of utilities
- 4) Value as a guide to stimulus sequencing in interactive data collection.

## References

- J. D. Davidson, "Forecasting Traffic on STOL," Operational Research Quarterly, 24(1973), 561-569.
- John A. Fiedler, "Condominium Design and Pricing: A Case Study in Consumer Trade-Off Analysis," Proceedings of the ACR (1972), 279-293.
- Paul E. Green and Michael T. Devita, "An Interaction Model of Consumer Utility," Journal of Consumer Research, 2(September, 1975), 146-153.
- Paul E. Green and Yoram Wind, Multiattribute Decisions in Marketing: A Measurement Approach (Hinsdale, Ill.: Dryden Press, 1973).
- Paul E. Green and Vithala R. Rao, "Conjoint Measurement for Quantifying Judgmental Data," Journal of Marketing Research, 8(August, 1971), 355-363.
- Richard M. Johnson, "Trade-Off Analysis: A Method for Quantifying Consumer Values," Unpublished Manuscript, Market Facts, Inc., Chicago (September, 1972).
- Richard M. Johnson, "Trade-Off Analysis of Consumer Values," Journal of Marketing Research, 11(May, 1974), 121-127.
- Richard M. Johnson, "A Simple Method for Pairwise Monotone Regression," Psychometrika, 40(June, 1975), 163-168.
- Richard M. Johnson, "On the Measurement of Utilities for the Direction of Public Policy," Unpublished Manuscript, Market Facts, Inc., Chicago (August, 1975).
- Richard M. Johnson and Gerald J. Van Dyk, "A Resistance Analogy for Efficiency of Paired Comparison Designs," Unpublished Manuscript, Market Facts, Inc., Chicago (April, 1975).
- Joseph B. Kruskal, "Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data," Journal of the Royal Statistical Society, Series B, 27(1965), 251-63.
- Richard B. Ross, "Measuring the Influence of Soft Variables on Travel Behavior," Traffic Quarterly (July, 1975), 333-346.
- J. Srinivasan and Alan D. Shocker, "Estimating the Weights for Multiple Attributes in a Composite Criterion Using Pairwise Judgments," Psychometrika, 38 (1973), 473-493.
- Dick Westwood, Tony Lunn, and David Beazley, "The Trade-Off Model and Its Extensions," Journal of the Market Research Society, 16(July, 1974), 227-241.

# ATTRIBUTE RATINGS AS PREDICTORS OF CLAIMED AND ACTUAL BEHAVIORS

Penny Baron and Gerald Eskin  
The University of Iowa

## Abstract

Cash register receipts from grocery shopping trips and interview data about store attributes (price, location and quality) and behavior concerning grocery stores shopped at were collected from 80 respondents over five weeks. Results indicate a high correspondence between claimed and actual behavior and similar but not identical regression translations between attribute scores and claimed and actual behaviors.

## Introduction

Attribute models of consumer choice assume that goods consist of attributes or characteristics which satisfy consumer needs. Goods vary in the attributes they contain or are perceived to contain and consumers choose among them based on their preferences for the attributes. Consideration of such models leads one to inquire after the attributes that consumers want in particular choice situations, their preference for them, the attribute content of the various goods available in the situation and the relationship between these factors and choice. But how best to obtain these assessments is problematic.

Empirically, use of marketplace data to inform multi-attribute models encounters serious problems. For example, variables of interest may not display adequate variation or relevant data may simply be unavailable. Observations are typically available only for aggregates of consumers. Thus, individual preference patterns are not revealed.

To directly assess individual consumers' preferences for attributes, most researchers turn to survey research or to laboratory experiments. In these settings, while data on individual attitudes and preferences can be obtained, marketplace behavior is not typically observable. Instead, one must settle for indicators of these behaviors, for example, overall attitude toward a product or respondents' claims about their behavior with respect to it.

Early applications of psychologists' work on attribute models focused on the tie between perceptions and evaluations of attributes and overall attitude or preference for the good in question. The presumption underlying this research was that such attitudes ultimately affected a consumer's marketplace choices. Subsequent work, however, has suggested that these relations, when they exist, can be expected to be quite weak. Overall attitude toward a good is only one of a number of factors affecting choices. Moreover, it was not quite clear in this early work how similar the relation between overall attitude and perceptions or evaluations of attributes was to the relation between these attribute responses and actual marketplace choices. Correspondence was often much less than perfect even under carefully controlled research situations in which respondent attitudes and behaviors were assessed under very similar conditions.

As a consequence, researchers have sought other variables which might tie beliefs and preferences concerning attributes more closely to actual behavior. Respondent claims about past and future behaviors consti-

tute an important class of such variables. Recently, much research has been directed toward investigating the relation between various claimed behaviors with respect to past or future marketplace choices and beliefs or preferences with respect to relevant attributes.

This research has tended to follow the same paradigms as those used earlier to study the overall attitude variable. It attempts to assess relationships under conditions which are most likely to produce a strong association. For example, researchers try to obtain measures of self-reported behaviors and observations of actual behaviors within the same situational context; they try to keep the time interval between the two assessments to a minimum; and they seek to examine behaviors whose execution is entirely under the control of the consumer. For example, in this spirit Wilson, Mathews and Harvey (1975) obtained a rather elaborate series of paper and pencil measures designed to assess behavioral intentions to choose brands of toothpaste and various motivational and cognitive determinants of those intentions. Immediately upon completing the required questionnaire, each respondent was given an opportunity to choose a toothpaste from among those about which he had just been questioned. Needless to say, under these circumstances, the relation between stated intentions and actual behavior was high. About 85 percent of the respondents selected the brand with the "highest ranked intention to purchase." Similarly, when researchers control situational factors and other intervening events, they find that people do what they intend to do. (See, for example, Darrock, 1971.)

While results such as these are comforting because they suggest that there is a high correspondence between claimed and actual behavior, the generality of this result remains problematic. "Actual" behaviors under survey or laboratory conditions like those described above may not be the same as "marketplace" behaviors. In such controlled research settings, constraints are either not present or the researcher attempts to minimize their effects. But in the marketplace, behavior is constrained. Thus, the ties between attitudes or beliefs concerning attributes and behavioral choices in controlled research settings may not be the same as those which obtain for behaviors in the "marketplace."

A related issue concerns the type of observations used in measuring choice behavior. Typically, only one, or at best, a few choices are observed and the observations are all in terms of the same operational definitions, e.g., product or brand chosen. However, in the marketplace, numerous definitions of choice behavior are possible, e.g., average amount spent on a given occasion, frequency of purchasing a product in a given month, total expenditure for a product in a given year, number of units of a product purchased in a given period, etc. How attributes are perceived and evaluated with respect to these various behaviors is not necessarily the same. In short, perceptions and evaluations of attributes may vary depending not only on whether one is concerned about claimed purchase behavior, choices in a research setting or actual marketplace choices, but also on the particular marketplace behaviors observed. The constraints operating on different marketplace behaviors might vary.

Several questions arise, then, about these relationships:

1. Do claims about behavior forecast actual marketplace choices?
2. Do attribute perceptions that correlate with claimed behaviors also correlate with marketplace choices?
3. How do differences in behavior patterns relate to differences in the role played by various attributes?
4. How are different types of marketplace behaviors related to attribute perceptions?

The present study investigates these questions in the context of a field study of store selection. In this study we followed the grocery shopping behavior of 80 respondents over a five-week period. The research procedures involved both field interviews and observation of actual marketplace behavior. During the interviews several types of paper and pencil measures on attributes, store perceptions, preferences, attitudes and claimed behavior patterns were obtained. Objective records of actual shopping behaviors were obtained by regular collection of all sales receipts accumulated by each respondent during the five-week period. These receipts were treated as a source of direct measures of actual marketplace behavior.

Below we describe the store attributes, the response measures, and the procedures used to obtain them. Analysis procedures for assessing the relationships among the attribute ratings and the response variables are compatible with linear compensatory attribute models. Interpretations and methodological implications are discussed in terms of the questions posed above.

#### Measurement Procedures

##### Sample

Respondents were selected using an area quota sampling procedure. Forty-three student interviewers were randomly assigned two 5-10 block sections of a small mid-western city. The sections spanned the entire city. Interviewers were instructed to obtain two eligible respondents using a standard door-to-door sampling procedure. An eligible respondent was a primary grocery shopper for a household which normally ate dinner at home.

##### Behavioral Measures

Out of the initial sample of 86, 6 withdrew from the study or failed to provide complete and usable data, e.g., they did not consistently complete records of their grocery shopping trips. Respondents who agreed to participate in the 5-week study filled out a record form for each grocery shopping trip they made and sealed all sales receipts from the trip in an envelope attached to the trip record form. These sales slips, together with corroborating information on the trip records about which stores were shopped at, constituted the source of measures on actual marketplace behavior.

The primary measures considered here are:

1. Store to which the most visits were made, the "most frequented" store.
2. Store at which the "most dollars" were spent.

##### Measures of Attributes and Perceptions

Interviews with respondents were conducted at the be-

ginning of the study and at the end of the 1st, 3rd, and 5th week. Measures of respondent perceptions, and identification of relevant attributes were based on the responses obtained during these interviews.

1. Relevant Attributes. During the 2nd interview contact, respondents were asked which store they liked shopping at best and their reasons for liking it better than others. Similar questions were asked about the store liked "2nd best," "least," store "shopped at most frequently" and stores "not shopped at." These responses were content analyzed. Six attributes were identified as most important based on "frequency of mention" across the respondent population. They were:

Prices - How costly the store's products are considering both regular prices and specials.

Location - The amount of time required to travel to the store.

Efficiency - How quickly products can be found in the store, checked-out and carried-away, for example, quick pick-up and easy parking.

Atmosphere - How pleasant it is to shop at the store, its appearance, cleanliness, and how friendly, helpful and courteous its personnel are.

Quality - Overall quality of the food sold.

Selection - The extent to which the store has the items wanted.

In the analyses reported here, only price, location and quality are considered. Price and location received much higher "frequency of mention" counts in the original content analysis than did the other 4 attributes. The smallest intercorrelations also obtained among price, location and the other 4 attributes. However, the intercorrelations among the remaining four attributes were relatively high. Of these, quality correlated least with the price and location attributes. Because it also seemed to be the most general of the four remaining attributes it was selected to represent them. Including the other three scales tends to make only marginal improvements in overall model fit. (Some comparisons of models specified in the three and six attributes are found in Section 1 of Appendix.) Factor analysis results were also consistent with this decision.

2. Perceived Attributes of Stores. Respondents were asked during the 3rd interview to rate each of the stores they had shopped at in terms of the above characteristics. They used a 5-point rating scale whose range of values went from "far above average" (5) to "far below average" (1). They were instructed to rate each store they were familiar with on each attribute in terms of their own impressions of the store. (See Section 2 of Appendix.)
3. Claimed Shopping Behaviors. On the 1st interview contact respondents were asked which store they shopped at most frequently and which they spent most of their food dollars at.

#### Analysis Procedures and Results\*

\* We thank Stephanie Takazawa for her extensive contribution to the data processing and analysis phase of this work.

## Procedures

For each respondent, perceived attributes of stores and claimed and actual behaviors were coded separately by store. The actual behavioral measures were recorded from the trip record forms for the collection of stores actually shopped at during the 5 weeks. Measures on attributes and claimed behaviors were obtained during the interviews for the collection of stores with which the respondent claimed familiarity.

These procedures resulted in 3 groups of respondent-store records. The first group consists of "familiar only" respondent store records. It contains those stores which the respondents claimed they were familiar with but which they did not shop at during the research period. The second group, a "shopped familiar" group, consisted of records for stores that the respondent claimed to be familiar with and at which he actually did shop at least once during the research period. These two groups of data are sometimes combined into an "all familiar" set, i.e., all stores the respondents claimed to be familiar with, including those they shopped at and those they did not. This "all familiar" set of store records includes "0" observations on the behavior measures--is store most frequented and is store where most \$ are spent?--for the stores where the respondent made no transactions. The third group of respondent-store records consists of stores which were shopped at during the study but with which respondents did not claim to be familiar. Thus no attribute ratings or attitude measures are available for this "shopped at only" group. Subsequent analysis did not include this set of data. (See Section 3 of Appendix for a discussion of the "shopped-at-only" respondent-store records.)

The principle analysis technique for the questions of interest here involved fitting multiple regression models which relate perceived store attributes to claimed and actual behavioral measures. The measure of perception of store attributes consists of respondents' ratings of the stores in terms of the price, location and quality attributes. Since the dependent variables can be treated as dichotomous (e.g., "most frequented store" and "all others") this approach is essentially similar to a discriminant analysis.

Implicit in this regression approach over the two previously defined data sets is the assumption that consumers possess identical and linear utility functions in the attributes considered but that they may differ in their perceptions of stores in terms of these attributes. It is clear in our case that these latter differences among consumers should exist. The best example is, of course, location. Differences in the relationship between the location of stores and of residence make it mandatory that attribute evaluations be considered a heterogeneous component. The homogeneity of the utility function across consumers may be questioned although previous empirical work has usually been unable to falsify it. (See, for example, Nakanishi and Bettman, 1974.) Here we maintain the assumption for simplicity. (In Section 4 of the Appendix evidence is given to support the assertion that this simplification does not substantially restrict the fit of the various models considered.)

### Relation Between Claimed and Actual Behavior

Are the claims that people make about their consumption patterns consistent with their actual behaviors? An important assumption of most attribute model research based on questionnaire or interview data is that they are consistent. Violation of this assumption could affect interpretation of the role of attribute perceptions

in determining marketplace choices. It may also affect interpretation of the role of consumers' attitudes toward their choice alternatives. Clearly, respondents' attitudes towards stores are predicated on their beliefs about what their actual shopping experiences have been. To the extent that what people do and what they claim to do are similar, variations between ties among marketplace behaviors and attributes and ties among claimed behaviors and attributes can be interpreted in terms of variations in the constraints surrounding each type of relation rather than in terms of respondent misperceptions of their past behavior or their inability to forecast future behavior.

The present study contains two sets of measures which enable a direct assessment of the relation between claimed and marketplace behavior. The measures are claimed vs. actual store which was "most frequented" and claimed vs. actual store at which "most food dollars" were spent. As indicated in Table 1, seventy-two respondents reported that they had a single store that they frequented most often. Based on actual counts of store visits, 63 (87.5%) frequented this store more than any other during the 5-week test period. Five other respondents claimed that they tended to shop at a set of two stores most frequently. Of this group 4 out of the 5 actually shopped at one of these stores more than any other. Seventy-three respondents claimed that they spent more of their food dollars at one particular store than at any other. Of this group, there were 69 (94.5%) matches between the claimed and actual modal expenditure store. Again, 5 respondents claimed that there were 2 stores in their modal set and 4 out of the 5 had actual modal values that matched their claim.

TABLE 1  
Relation Between Claimed and Actual Behavior

	Actual Shopping Behaviors		
	Total	Same	Different
Single store claimed "most frequented"	72	63 (87.5%)	9 (12.5%)
Two stores claimed "most frequented"	5	4 (80.0%)	1 (20.0%)
Single store claimed "most \$" spent	73	69 (94.5%)	4 (5.5%)
Two stores claimed "most \$" spent	5	4 (80.0%)	1 (20.0%)

These results suggest that there is reasonable consistency between these two types of measures even though the actual behaviors occurred subsequent to obtaining the claimed measures. Thus, these latter measures constituted a forecast of a future behavior pattern rather than a description of past behaviors. As such, their consistency is encouraging since the conditions surrounding the actual behaviors were those the respondents actually encountered in choosing stores to shop at during the five-week study period.

### Relation Between Store Attributes and Claimed and Actual Behaviors

Here we consider the identification and measurement of the relationship between attribute ratings and the measures on store "most frequented" and store spent "most dollars" at. Table 2 compares regression coefficients and t statistics for models which relate the price, location and quality measures to the dichotomous frequency and expenditure measures. Results are shown separately for the claimed and actual measures of behavior and for



the "shopped familiar" and "all familiar" sets of store records.

TABLE 2  
Relation of Store Attribute Ratings to Claimed and Actual Store "Most Frequented" and Store at Which Spent "Most Dollars"  
 (Multiple regression coefficients  $\beta$  and t statistics)

Most Frequented	"Shopped Familiar" Sample (N = 216)		"All Familiar" Sample (N = 390)	
	Actual	Claimed	Actual	Claimed
Price	.140(4.3)	.237(5.8)	.110(5.4)	.176(6.9)
Location	.132(4.6)	.092(2.6)	.122(7.5)	.111(5.6)
Quality	-.015(0.3)	-.004(0.0)	.022(0.9)	.036(1.3)
R <sup>2</sup>	.16	.18	.20	.21
Most Dollars				
Price	.183(5.8)	.205(6.3)	.134(6.8)	.152(7.4)
Location	.115(4.1)	.094(3.3)	.109(7.0)	.104(6.4)
Quality	-.020(0.6)	-.001(0.0)	.018(0.8)	.033(1.4)
R <sup>2</sup>	.20	.21	.23	.24

Comparison of the two measures of claimed behavior and the two measures of actual behaviors enables an assessment of the extent to which they are similarly related to perceived attributes of the stores. Comparisons are also possible with respect to the type of behavior at issue, frequency of visiting a store or amount spent at it.

With respect to the "most dollars" measure, good matches are obtained in both data sets.

Coefficient values are similar and the attribute rank orders are identical--price first, location second and quality last. The "t" statistics rank in a similar fashion. All tests on price and location are significant ( $p < .01$ ) but those for quality are not. The only differences here between the results for actual and claimed behavior are that location is relatively more important than price in the translation to actual behavior than in the translation to claimed behavior.

For the "most frequented" measure, there is also some agreement across dependent measures but less than in the case of the dollar measure. Price and location again emerge as the terms with large and significant coefficients but their internal ranking is less similar. The results for claimed frequency indicate that price is the most important factor while the results for actual frequency indicate a stronger role for location. In both the "all familiar" and "shopped familiar" data sets, the regression coefficient of the price attribute is substantially lower for the actual behavior measure than for the claimed behavior measure. The t statistic parallels this shift. The coefficient for location is somewhat larger for actual than for claimed behavior and it becomes much more significant. Differences in results across samples appear slight for these data. Consideration of stores "not shopped at" does not seem to substantially alter the role of perceived attributes for any of the four measures.

One interpretation of these results is that consumers do not select a store for shopping trips on which they plan to spend a lot unless they perceive prices there to be good. For minor, convenience type trips which involve small expenditures, location increases in relative importance. This would explain the shift in coefficient

size between the measures of actual "most frequented" and actual "most dollars." The higher coefficients of price for the "claimed" frequency measure may be attributed to consumers not making a clear separation between frequency and amount of expenditures when responding to the "claimed" question. It is also possible that they are unaware of their own choice patterns which appear to give more importance to convenient location and the related conservation of valuable time for minor shopping trips but to give them less weight when shopping trips involving large expenditures are being undertaken.

#### Conclusions

These data suggest that strong relationships do exist between what people claim they do and what they actually do. However, there is not a one-to-one correspondence, even for this relatively simple and important behavior, choice of a grocery store to shop at. While recall of such behavior is very good, and we suspect much better than for behaviors like "brand of chocolate bar last purchased" or "quantity of sugar consumed in a typical month," ambiguities do creep in.

The translation from attributes to claimed and actual behavior measures are similar but not identical. Differences in the consistency of results between claimed and actual store "most frequented" and claimed and actual store at which "most \$ were spent," may be related to the differential clarity of the notion of "most frequented" as opposed to "most \$." Asking about expenditures is a clearer question in the sense that it elicits consideration of shopping trips involving significant expenditures as opposed to "all trips" suggested by the "most frequented" measure.

In the absence of a clear understanding of how consumers process their perceptions of stores, all such conclusions must remain tentative.

#### Appendix

- Columns 1 and 2, Table 3, compare the R<sup>2</sup> coefficients for regression models specified to include all six attributes and the sub-set of 3 [prices, location and quality] considered in the body of the paper. Differences in the fit of these two models are small. However, the marginal contribution of the three additional attributes in reducing explained variance is statistically significant ( $p < .05$ ) for most response measures. Thus it would be erroneous to conclude that only 3 attributes are operating here.

TABLE 3  
R<sup>2</sup> Values under Different Model Specifications for "All Familiar" Sample (N = 390)

# of Attributes:	Model 1	Model 2	Model 3	Model 4
Weighting: none	6	3	3	3
		none	Absolute importance weights	Relative importance weights
Claimed "most frequented"	.24	.21	.21	.22
Claimed "most \$ spent"	.28	.24	.25	.25
Actual "most frequented"	.27	.21	.22	.22
Actual "most \$ spent"	.28	.23	.25	.24

2. Each respondent was also asked to rate each of these stores, using the same scale, in terms of his overall evaluation of it and to rate each store attribute in terms of its importance to him in choosing a grocery store to shop at. Respondents were also asked in the second interview to tell which store they "liked best" to shop at and which they liked to shop at "least." However, we do not treat these variables here.
3. About 50 percent of the respondents claimed to be familiar with each and every store that they actually shopped at. The remaining 50 percent failed to claim familiarity with about 40 percent of the stores they actually shopped at. Overall, then, for about 20 percent of the cases in which a respondent actually shopped at a store, he did not claim to be familiar with it. Visits to these stores accounted for about 13 percent of all visits made during the study period and for about 10 percent of the total dollar volume. These figures suggest that these stores are visited infrequently and that relatively few food dollars are spent at them.

Some portion of the respondents' failure to claim familiarity with one particular store may have been due to its closing during the research period. Some respondents visited this store only once. This was for its close-out sale.

4. With respect to differences among consumers in their utilities for the attributes, the question here is whether allowing for such differences improves the predictive value of the store attribute ratings. This was investigated by comparing 3 alternative procedures for weighting the relative contributions of each store-attribute rating. The first method employed standard regression techniques to produce the weights. The standard regression coefficients are estimates of the relative importance of the attributes, assuming that all consumers have similar utility functions. The 2nd procedure involved weighting the store attribute ratings of each consumer by his self-reported importance ratings. In this version, store-attribute ratings were first transformed so that the scale values ranged from -2 to +2 rather than from 1 to 5, the average or "normal" level of an attribute now being coded "0" rather than "3." These attribute scores were then multiplied by their absolute importance ratings before subjecting them to regression analysis. The 3rd method utilized relative importance scores to weight the transformed store-attribute ratings. Columns 3 and 4 of Table 3 present the  $R^2$ 's obtained using each of these methods. Consistent with previous work on this issue, neither of the two procedures, which used weights based on individual consumer's importance ratings of the attributes, resulted in more explained variation than the standard regression procedure. The magnitudes of the regression coefficients were also similar across the three procedures.

#### References

R. I. Darroch, "Attitudinal Variables and Perceived Group Norms as Predictors of Behavioral Intentions and Behaviors in the Signing of Photographic Releases." Unpublished doctoral dissertation, University of Illinois, 1970.

M. Nakanish and J. R. Bettman, "Attitude Models Revisited: An Individual Level Analysis," Journal of Consumer Research, 1(1974), 16-21.

D. T. Wilson, H. L. Mathews and J. W. Harvey, "An Empirical Test of the Fishbein Behavioral Intention Model," Journal of Consumer Research, 1(1975), 39-48.

INFLUENCE PROCESSES IN INTERPERSONAL PERSUASION

Morris Holbrook, Columbia University<sup>1</sup>  
 John O'Shaughnessy, Columbia University

Abstract

This paper reviews the progress of research on interpersonal persuasion in the salesman-buyer dyad. The nature of the dynamic verbal interactions between salesmen and buyers is identified as a promising focus for research. Methodological suggestions are therefore offered for designing content analyses to represent the unfolding of salesman-buyer interactions over time.

Introduction

The verbal interactions between a salesman and buyer engaged in the selling process are a singularly clear example of interpersonal influence at work. This paper briefly outlines some key distinctions between various ways of studying interpersonal influence in the salesman-buyer dyad. It then employs these distinctions to analyze the development of hypotheses and methodologies directed at studying the selling process. Finally, it draws conclusions concerning the approach of greatest potential use in future studies of interaction between salesmen and buyers.

A General View of Selling Interactions

Robertson and Chase (1968) portrayed the selling process as an open dynamic system in which environmental inputs (e.g., human, technological, and organizational) are transformed into outputs (e.g., cognitive, affective, and behavioral) which are, in turn, fed back into subsequent stages in the salesman-buyer dyad (e.g., as different expectations, feelings, skills, or resources). The heart of the system is the transformation process which converts inputs into outputs and which involves the essence of salesman-buyer interaction: the similarity between participants, their perceptions of one another, how fully they fulfill reciprocal role expectations, and most especially, how they talk and react to each other. It is this interaction process which lends the salesman-buyer dyad its dynamic character.

Unfortunately, the almost self-evident importance of the verbal interaction between salesman and buyer has inspired surprisingly little serious empirical research--virtually none of which, we shall argue, deals adequately with the issues involved. Most of the relatively few well-known studies were reviewed by Davis and Silk (1972) and by Capon, Holbrook, and Hulbert (undated). But neither review adopted a critical perspective sharply focused on revealing the most fruitful direction for future research in this area. For example, Davis and Silk distinguished between three types of salesman-effectiveness research: "psychological tests" (e.g., selection studies), "interpersonal influence" studies (e.g., based on mass-communication theory), and "customer-salesman interaction" studies (e.g., focusing on the similarity or matching between buyer and salesman). There is, however, a conceptually distinct fourth type of selling-process research which

these authors neglected, but which may offer the greatest potential for explaining the outcome of selling interactions.

A Typology of Research on the Selling Process

Research on the selling process has been characterized by two key distinctions.<sup>2</sup> First, such studies have typically focused either on static characteristics of the buyer and salesman (e.g., demographic, socioeconomic, or personality variables) or on communication variables (e.g., those represented by the familiar communication-theoretic paradigm--source-message-channel-receiver-effect). An especially important communication variable is the verbal content of the selling interaction. Secondly, research may view the selling process as a two-way interaction (focusing simultaneously on the actions and reactions between buyer and salesman) or may adopt an essentially one-way perspective (dealing separately with the salesman's and buyer's characteristics and/or behavior). These contrasts suggest the typology of selling research shown in Table 1, which assigns each cell a convenient name (and indicates Davis and Silk's terminology parenthetically).

TABLE 1  
 Typology of Research on the Personal Selling Process

	Focus on static characteristics	Focus on communication variables
one-way perspective	selection studies (psychological tests)	communication model (interpersonal influence)
two-way perspective	similarity or matching studies (customer-salesman interaction)	interaction synchronization (-----)

The Progress of Research on the Selling Process

Selection Studies

Studies of salesman selection (e.g., Mayer & Greenberg, 1964) have been ably reviewed by Cotham (1970) and will not concern us here, except to point out that sometimes studies which purport to deal with salesman-buyer interaction may actually confine their attention to a rather limited, one-way analysis of static salesman

<sup>1</sup>Morris Holbrook gratefully acknowledges the support of Columbia University's Faculty Research Fund.

<sup>2</sup>Please note that these distinctions are not synonymous with those introduced by Capon et al. (undated).

characteristics. Albaum (1967), for example, predicted the sales penetration of service stations using independent variables like the dealer's business attitudes, number of services provided, and customer waiting-time. But clearly, such factors do not adequately represent the complex, interactive verbal communication between buyer and salesman.

#### Communication Model

The effects of both the sales message and its source have received attention from researchers approaching selling research from a communication-theoretic perspective.

Message effects. In an early study of communication in selling, Pace (1962) found a significant tendency for higher-performing door-to-door saleswomen to earn higher interview-based ratings on their use of emotional appeals, dramatization, and language (but not on several other communication characteristics). More directly, Levitt's (1965) well-known experimental study of one-way persuasive communication compared the effects of contrasting sales messages themselves, demonstrating the significant superiority of "good" over "poor" filmed presentations in persuading subjects to react favorably to a new paint product.

Such clearcut results, however, have often failed to appear in subsequent studies of message effects. Farley and Swinth (1967), for example, found that a "product pitch" (describing the features of a roll-up yardstick) and a "personal pitch" (relating the yardstick to the subject's role as student or wife) exerted significantly different effects on attitude, but not on choice behavior. Similarly, though Reizenstein's (1971) hard- and soft-sell messages were equally ineffective in causing reversal of preference between first- and sixth-ranked patterns of crystal stemware, the soft-sell did produce significantly greater decision reversal between the first- and second-ranked alternatives, a difference which Reizenstein explained by invoking the theory of cognitive dissonance.

Even more confusing findings have emerged from a recent series of communication-theoretic studies comparing "canned" presentations with "interactive" deliveries in which the salesman encourages the prospect to talk. Jolson's (1975) interactive presentations of a new reference book (in which a classroom audience was permitted to ask questions) generated greater product comprehension, but his canned version (with heavy use of audiovisual aids) produced a higher rate of purchase intention. It appears dangerous, however, to generalize even this equivocal finding from Jolson's classroom setting to a selling situation involving one-to-one interaction.

Capon (1975) avoided this methodological weakness by comparing the effectiveness of canned and interactive sales presentations in the telephone selling of magazine subscriptions. Using six messages in a complex quasi-factorial design, he found no significant main effect of his canned/interactive manipulation on attitudes or purchase behavior, though significant interaction effects did suggest that certain salesmen were most successful with certain messages.

Another possible situational determinant of communication effectiveness is the business environment characterizing the seller's product. Pasold (1975) hypothesized that selling strategies which he called "preactive" (salesman dominated), "interactive" (mutually directed), and "proactive" (focusing on the alignment of long-term objectives) might, respectively, be most effective in markets characterized by imperfect competition, oligopoly, and turbulence. A correlational study of stocking sales found beta coefficients with the pre-

dicted relative magnitudes for the contrasting selling strategies. But the tiny  $t$  values of these coefficients and the failure of their relative magnitudes to be consistently upheld in a validation sample raise questions concerning the statistical significance of Pasold's finding.

Source effects. Levitt's (1965) pathbreaking study also extended the Hovlandian tradition of one-way communication-theoretic research to an examination of the effects of source credibility in selling. He claimed to find both a source effect (more favorable response to the salesman identified with the more credible parent firm) and a sleeper effect (the erosion of this differential over time), but Capon, Holbrook, and Hulbert (1972) have questioned the significance and consistency of these findings, without serious challenge from Levitt (1972).

Conclusions on the communication model. Perhaps the safest conclusion to be derived from the one-way communication-theoretic studies is that the nature of source or message effects tends to depend upon the situation--on the dissonance of the customer or his social setting or his profession, on the salesman's personal characteristics, on market conditions--or, in short, on the two-way process of interaction between buyer and salesman. Apparently, a research paradigm which views the selling process as one-way communication is inadequate to the task of generating consistent findings. Let us turn, then, to studies which have viewed selling as a two-way process of interaction.

#### Similarity or Matching Studies

In search of similarity effects. Attempts to relate sales success to the similarity or matching of salesman and prospect on static characteristics have proven inconclusive, almost without exception. Evans' (1963) classic study in this area claimed such relationships in the sale of life insurance, but when reanalyzed by Capon, Holbrook, and Hulbert (undated), his data did not pass conventional tests of statistical significance. Gadel (1964) showed that life-insurance buyers tended to resemble salesmen in age, but because she looked only at purchasers, no inference can be made concerning the effects of age similarity on sales success. Capon (1975) found no significant relationship between the perceived similarity of a magazine salesman and attitudes or intention to subscribe. Mathews, Wilson, and Moncky (1972) manipulated the perceived similarity/dissimilarity of students engaged in a selling simulation and found the expected difference between treatment groups in cooperative and competitive behavior, but interpretation of these results was beclouded by a still higher incidence of cooperation and conflict in a control group than in the respective treatment groups. Finally, Tosi (1966) adopted a slightly different view of the similarity hypothesis, focusing on the matching of role expectations between buyers, salesmen, and sales managers; but with various measures of sales success, only one of his seven hypotheses received statistically significant support.

Comparisons of similarity and source effects. Two ambitious studies have pitted the already shaky source and similarity hypotheses against one another, again with mixed results. Brock's (1965) two confederates tried to switch paint customers to a different-priced brand using either a similarity or an expertise appeal (claiming recently to have purchased about the same amount of paint or twenty times as much). The results supported the similarity hypothesis at the expense of the source effect, but their interpretation was complicated by the questionable believability of the expertise manipulation, the salesman's apparent complicity in the experimental design, and the failure of that design to permit

the simultaneous support of both hypotheses. Indeed, when Woodside and Davenport (1974) created a factorial manipulation of both perceived similarity (a record-store saleswoman's claimed preference for the same or a different type of music) and perceived expertise (her apparent ability to explain the proper use of a tape-recorder cleaning device), both treatment variables exerted significant positive main effects on purchase of the cleaning gadget, with a significant difference in favor of the source effect. Perhaps these results differed from Brock's because the salesgirl convincingly demonstrated her expertise by explaining the product's use whereas the paint salesman only claimed to have bought an almost preposterously large amount of paint.

Conclusions on source and similarity effects. Clearly, the debate between advocates of source and similarity effects could be continued, probably without ever resolving the issue definitively. It appears, however, that a clearer understanding of selling might emerge from a reorientation of research on the salesman-buyer dyad toward a more deliberate focus both on the two-way interaction between buyer and salesman and on the verbal content of that interpersonal communication process. This is the approach we call the study of "interaction synchronization" and to which we now turn.

The Study of the Synchronization of Verbal Interaction

Unfortunately, we cannot point to examples of research which has successfully viewed selling as a dynamic process of verbal interaction. We can, however, cite several studies which have helped to clear a path to examining the synchronization of verbal behavior.

Willett and Pennington (1966) were the first to focus systematically on a content analysis of the verbal components of buyer-salesman interaction. They applied Bales' (1950) Interaction Process Analysis to protocols of the transactions between 14 appliance salesmen and their customers and found, for example, that successful sales outcomes were associated with a high incidence of giving suggestions and asking for opinions and with a relatively higher involvement of salesmen in asking task-oriented questions. Pennington's (1968) own content analysis of the same data viewed the salesman-buyer dyad as an example of bargaining behavior. A discriminant analysis successfully predicted about 80% of the purchase outcomes using the frequencies of verbal acts falling into various bargaining categories (direct offers, attempts to change concession limits, commitment to concession limits, etc.). A further re-analysis of the same protocols by Olshavsky (1973) counted the number of verbal acts referring to various product-related attributes during succeeding stages of the sales interview and highlighted, for example, the customer's predominant concern for price during what he called the "evaluation" phase and the relegation of concern for nonproduct attributes like delivery or service to the concluding "consummation" phase. Thus, these three studies have quite appropriately focused attention on various facets of the verbal content of selling interactions. They have, however, shared a common tendency to treat these verbal interactions in an aggregative fashion (e.g., by summing the frequency of verbal acts in each category across customers and/or salesmen) with no explicit attempt to characterize the interaction through time of the content of the verbal acts within a specific customer-salesman pair. Because of this failure to investigate the unfolding interaction of the salesman's and buyer's verbal behavior, these studies tell us little about how a salesman might adjust the verbal content of his presentation to the responses of the buyer in order best to achieve sales success.

Ironically, the oft-forgotten studies by Chapple (Chapple & Donald, 1947; Norman, 1954) displayed just the opposite virtue, focusing resolutely on the temporal relationship between the speaking (or gesturing) of a salesman and that of his interlocutor. Chapple's Interaction Chronograph generated such indices as the salesman's activity level (the average length of his verbal and nonverbal acts) and his initiative-dominance level (his relative tendency to break a silence and to keep talking when interrupted). Work with such measures found good fits between the sales-performance rank of salesmen and the degree to which their interaction-timing indices resembled those of highest-ranking performers. Yet, while elegantly handling the temporal pattern of communication behavior, Chapple's method almost completely ignored the verbal content of salesman-buyer interactions. The Interaction Chronograph might help us determine when a salesman should talk, but it cannot tell us what he should say.

An Approach to Studying the Synchronization of Verbal Interaction in Selling

An Illustration of the Content Analysis of Interaction Synchronization

Apparently what we need, then, is research relating selling success to the way in which the salesman adjusts both the timing and the content of his verbal behavior to that of his customer. The raw materials for such a methodology would include, first, some content-analytic scheme theoretically derived from the nature of interpersonal influence and, secondly, some way of analyzing the temporal unfolding of verbal behavior.

Let us begin by presenting a rather idealized illustration of such an approach to the content analysis of interaction synchronization. Elaborating upon O'Shaughnessy's (1971/72, 1974) analysis of interpersonal influence, the illustrative content-analytic scheme shown in Table 2 was developed by arraying Kelman's (1961) types of influence (compliance, identification, internalization) against two of Etzioni's (1964) means to power (material or tangible and symbolic or intangible).

TABLE 2  
Classification of the Verbal Content of Salesman-Buyer Interactions (With Examples of Each Type of Interaction)

Kelman's types of influence (source of influence)	Etzioni's Means to Power	
	a. Material means (tangible)	b. Symbolic means (intangible)
1. compliance (means control)	1a. patronage	1b. approval or agreement
2. identification (attractiveness)	2a. physical contact (nonverbal)	2b. reference to role-related norms or social values
3. internalization (credibility)	3a. factual information	3b. evaluative information

As suggested by Bales (1950), each verbal act by the buyer (B) or salesman (SM) might involve one or more instances of seeking (S) or offering (O) a particular type of content. The protocol of an illustrative salesman-buyer dialogue might then be coded interaction-by-interaction as follows:

1. SM-O-1b, SM-S-2b, SM-O-3a (salesman offers approval, seeks information concerning the buyer's values, offers factual information concerning a tangible feature of the product)
2. B-O-2b, B-S-3b (buyer answers question about values and seeks evaluative information about an intangible feature)
3. SM-O-2a, SM-O-3a (salesman grabs buyer by the arm and, instead, tells him something about a tangible feature)
4. B-O-1b, B-S-3b (buyer offers approval and repeats his question)
5. . . . .

Such a content-analytic coding of the interactions between buyer and salesman could generate a measure of the degree of match or synchronization between their verbal behavior--that is, the extent to which each participant offers the type of verbal acts which the other seeks. Thus, for example, interaction synchronization might be defined as the relative frequency with which verbal acts of type B-S-i are followed by verbal acts of type SM-O-i ( $i = 1a, \dots, 3b$ ). Alternatively, the concept might be operationally defined as the average proportion of content categories represented by each verbal act which have corresponding content components in the preceding verbal act. The latter measure would ignore whether information was sought or offered, but would represent focusing on the same topic. Obviously, a myriad of possible content-matching measures exists, given the basic content data, with little a priori basis for choosing between them. The crucial point is that such measures would depict some facet of interaction synchronization as a basis for predicting sales success.

#### Methodological Problems in Content Analysis

The most obvious objection to the kind of content-analytic methodology suggested above is its enormous difficulty of execution. Whatever the content-analytic classification dictated by the researcher's theory of interpersonal influence, perplexing methodological issues arise, first, in defining the content unit to be categorized (a word, a phrase, a sentence, a complete thought, the verbal act, or even the entire transaction) and, secondly, in formulating operational definitions of each content category which are clear, tight, and valid enough to permit some degree of inter-coder reliability. Certainly, there are other tough methodological issues to be solved in research on verbal interaction in selling, but exploratory attempts in this direction at Columbia have convinced us that they are dwarfed by the problems of the content analysis itself (O'Shaughnessy, 1971/72). The difficulty of a research technique, however, has never been a persuasive argument against its potential usefulness, and we remain convinced that refinement of the investigation of personal influence in selling depends upon our ability to deal meaningfully with the interaction between the content of the verbal behavior of buyer and salesman. Even though we may be using rough tools, they appear to stand a better chance of reaching the heart of the matter.

#### Simplifying the Content-Analytic Scheme

Accordingly, we have been searching for ways to simplify the rather idealized content analysis outlined above. One solution is based on the assumption, clearly articulated by Winch (1958) that human action is rule-governed behavior and that these rules are reflected in the concepts used by the individual to describe his actions. If the salesman is to understand how a buyer is likely to act, he must familiarize himself with the concepts in terms of which the buyer is viewing the situation and recognize that these concepts reflect the rules being used to guide buying action. If these concepts change (e.g., through salesman/buyer interaction), so do the buyer's perceptions of the situation; his actions change as a result. Of course, salesmen are intuitively aware that words reflect attitudes, as when they speak of the buyer making an "investment" rather than a "payment" or when their product is described as "cheap" instead of "inexpensive." However, less obvious is the claim that concepts mirror the set of rules that direct a buyer on what action to take.

Since the notion of following a rule is inseparably linked to the possibility of making a mistake, Winch (like Wittgenstein) has defined behavior as rule-following whenever it makes sense to distinguish between a right and wrong way to do something. It follows that one practical guide to a buyer's rules is to ask him what errors he particularly seeks to avoid. Although this viewpoint has been the subject of much discussion and criticism, its utility for our purposes has not been seriously undermined (Gellner, 1971; MacIntyre, 1971). The immediate problem, however, is to relate this perspective to our own focus on interaction synchronization.

#### Hypotheses on Rule-Following in the Salesman-Buyer Dyad

Peters (1958) has argued that any satisfactory explanation of human action must constitute a "sufficient reason" in the sense that the action must be perceived as both an efficient and a socially appropriate means to goal attainment. From our viewpoint, then, a buyer would have (a) rules reflecting beliefs as to what constitutes an efficient means (offering) for attaining the goals stemming from his functional requirements and/or role responsibilities and (b) rules reflecting what he considers socially appropriate behavior on the part of the salesman.

If buying action is rule-following, tentative hypotheses are that the verbal interactions of more successful salesmen (1) will seek to identify the buyer's rules; (2) will have content that is relevant to a rule that has just been reflected in the concepts used by the buyer; (3) will have content that satisfies rules regarding socially appropriate salesman behavior; (4) will have content (if possible) that satisfies product-related rules or (if not possible) that provides the buyer with additional concepts to achieve modification of these rules.

#### Revised Categories for Content Analysis

The two types of rules suggested above (efficiency and social appropriateness) provide the broad outline for a content-analytic scheme congruent with the above discussion of the interpersonal influence process, except that it ignores what Kelman (1961) referred to as "compliance" as being of little relevance in selling. Such a classification scheme might include more detailed categories, as listed below.

1. Rules reflecting expectations of socially appropriate behavior on the part of salesmen. If the salesman ad-

heres to the buyer's rules for socially appropriate behavior, the buyer might tend to accept influence via Kelman's process of identification, dependent upon the salesman's "attractiveness." Such rules are associated with what we might call "the four I's": (a) Image Projected--the image the buyer wishes to project (e.g., a verbal focus on "saving money" might indicate the rule that "communications from the salesman must support an image of myself as a hardheaded businessman"); (b) Integrity Sought--the professional and personal norms he values (e.g., a statement that "firms who divulge a supplier's competitive bid to the competition are just not playing fair" might reflect the rule that "communications from the salesman must not contain inquiries about competitive bids and preferably should show respect for my objections to such breaches of confidence"); (c) Interpersonal Relations--how the salesman fits into the buyer's overall role set (e.g., remarking on "the need to confer with everyone affected by a purchase" might indicate the rule that "communications from the salesman must show a sensitive regard for the fact that I need to consult with others before giving an order"); (d) Innovation Risk--fear of the possible consequences of a departure from customary buying patterns (e.g., a reminder that "the firm has been using the same material for twenty years" might reflect the rule that "communications from the salesman must address the problems I face in acting as a change agent within my own organization").

2. Rules covering beliefs as to efficient means of achieving functional requirements and/or role responsibilities. If a salesman adheres to (or succeeds in modifying) the rules regarding beliefs about efficient means, the buyer might move toward influence acceptance via Kelman's "internalization," as dependent on the salesman's "credibility." Such rules relate to considerations which we might term "the four S's": (a) Suitable Performance/Quality Characteristics of the product and/or vendor (e.g., verbal reference to "the supplier's quality control" might indicate the rule that "communications from the salesman must satisfy me that his firm has the experience to manufacture at the quality level required"); (b) Sources of Supply--the likely availability of the vendor's product (e.g., statements foreboding "the ill effects of late delivery" might reflect the rule that "communications from the salesman must demonstrate his firm's ability to meet its promised delivery date"); (c) Size of Delivered Cost--as opposed to quoted price (e.g., reference to "the costs of handling a raw material" might indicate the rule that "communications from the salesman must show that total use costs are at least competitive"); (d) Services Expected (e.g., mention of "installation and training problems" might reflect the rule that "communications from the salesman must explain the extent of these problems and the help we would get from his firm").

#### A Simplified Measure of Interaction Synchronization

To repeat, the simplified content analysis which we envision would not attempt to categorize all verbal acts between buyer and salesman, but rather would focus only on that verbal behavior which can readily be classified into those rule-related categories outlined above. More specifically, each implicit rule expression by the buyer would be identified and the salesman's verbal behavior scored on whether or not it responds directly to the rule represented by the buyer's most recent verbal act--either by providing information congruent with the rule or by attempting to modify the rule itself by questioning the buyer in a way that encourages him to re-examine his assumptions while providing him with additional concepts. In other words, the protocol of a salesman-buyer transaction would be scored on the degree to which the verbal behavior of the buyer and salesman is linked by their common focus

on an implicit set of buying rules. In that sense, the measure would represent the salesman's responsiveness in offering the information which the buyer seeks and would therefore serve as an index of interaction synchronization as defined above. It is our hope that such a measure would predict sales success better than other currently available techniques.

#### References

- Gerald Albaum, "Exploring Interaction in a Marketing Situation," Journal of Marketing Research, 4 (May 1967), 168-72.
- Robert F. Bales, Interaction Process Analysis (Cambridge, Mass.: Addison-Wesley, 1950).
- Timothy C. Brock, "Communicator-Recipient Similarity and Decision Change," Journal of Personality and Social Psychology, 1 (June 1965), 650-54.
- Noel Capon, "Persuasive Effects of Sales Messages Developed from Interaction Process Analysis," Journal of Applied Psychology, 60 (1975), 238-44.
- Noel Capon, Morris B. Holbrook, and James Hulbert, "Industrial Purchasing Behavior: A Reappraisal," Journal of Business Administration, 4 (Fall 1972), 69-77.
- Noel Capon, Morris Holbrook, and James Hulbert, "The Selling Process: A Review of Research," unpublished paper, Columbia University, undated.
- Eliot D. Chapple and Gordon Donald, "An Evaluation of Department Store Sales People by the Interaction Chronograph," Journal of Marketing, 12 (October 1947), 173-85.
- James C. Cotham III, "Selecting Salesmen: Approaches and Problems," MSU Business Topics, 18 (Winter 1970), 64-72.
- Harry L. Davis and Alvin J. Silk, "Interaction and Influence Processes in Personal Selling," Sloan Management Review, 13 (Winter 1972), 54-76.
- Amitai Etzioni, Modern Organizations (Englewood Cliffs, N.J.: Prentice-Hall, 1964).
- Franklin B. Evans, "Selling as a Dyadic Relationship: A New Approach," American Behavioral Scientist, 6 (May 1963), 76-9.
- John U. Farley and Robert L. Swinth, "Effects of Choice and Sales Message on Customer-Salesman Interaction," Journal of Applied Psychology, 51 (April 1967), 107-10.
- M. S. Gadel, "Concentration by Salesmen on Congenial Prospects," Journal of Marketing, 28 (April 1964), 64-6.
- Ernest Gellner, "Concepts and Society," in Bryan R. Wilson, ed., Rationality (New York: Harper Torchbook, 1971).
- Marvin A. Jolson, "The Underestimated Potential of the Canned Sales Presentation," Journal of Marketing, 39 (January 1975), 75-8.
- Herbert C. Kelman, "Processes of Opinion Change," Public Opinion Quarterly, 25 (Spring 1961), 57-78.
- Theodore Levitt, Industrial Purchasing Behavior (Boston: Harvard University Graduate School of Business Administration, 1965).
- Theodore Levitt, "Industrial Purchasing Behavior: A Bayesian Reanalysis," Journal of Business Administration, 4 (Fall 1972), 79-81.

- Alasdair MacIntyre, "The Idea of a Social Science," in Bryan R. Wilson, ed., Rationality (New York: Harper Torchbook, 1971).
- H. Lee Mathews, David T. Wilson, and John F. Moncky, Jr., "Bargaining Behavior in a Buyer-Seller Dyad," Journal of Marketing Research, 9 (February 1972), 103-5.
- David Mayer and H. M. Greenberg, "What Makes a Good Salesman," Harvard Business Review, 42 (July-August 1964), 119-25.
- J. T. Norman, "Now They Stand Up Under Pressure," Sales Management, 43 (August 1, 1954), 36-7, 84-5.
- Richard W. Olshavsky, "Customer-Salesman Interaction in Appliance Retailing," Journal of Marketing Research, 10 (May 1973), 208-12.
- John O'Shaughnessy, "Selling As an Interpersonal Process," Journal of Retailing, 47 (Winter 1971-1972), 32-46.
- John O'Shaughnessy, "The Influence of Personal Communication on a Buyer's Choice Criteria in the Salesman/Buyer Dyad," in Jagdish N. Sheth, ed., Models of Buyer Behavior (New York: Harper & Row, 1974).
- R. Wayne Pace, "Oral Communication and Sales Effectiveness," Journal of Applied Psychology, 46 (October 1962), 321-4.
- Peter W. Pasold, "The Effectiveness of Various Modes of Sales Behavior in Different Markets," Journal of Marketing Research, 12 (May 1975), 171-6.
- Allan L. Pennington, "Customer-Salesman Bargaining Behavior in Retail Transactions," Journal of Marketing Research, 5 (August 1968), 255-62.
- R. S. Peters, The Concept of Motivation (London: Routledge and Kegan Paul, 1958).
- Richard C. Reizenstein, "A Dissonance Approach to Measuring the Effectiveness of Two Personal Selling Techniques Through Decision Reversal," Proceedings, American Marketing Association, Fall Conference, 1971, 176-80.
- Thomas S. Robertson and Richard B. Chase, "The Sales Process: An Open Systems Approach," MSU Business Topics, 16 (Autumn 1968), 45-52.
- Henry L. Tosi, "The Effects of Expectation Levels and Role Consensus on the Buyer-Seller Dyad," Journal of Business, 29 (October 1966), 516-29.
- Ronald P. Willett and Allan Pennington, "Customer and Salesman: The Anatomy of Choice in a Retail Setting," Proceedings, American Marketing Association, Fall Conference, 1966, 598-616.
- P. G. Winch, The Idea of a Social Science (London: Routledge and Kegan Paul, 1958).
- Arch G. Woodside and J. William Davenport, Jr., "The Effect of Salesman Similarity and Expertise on Consumer Purchasing Behavior," Journal of Marketing Research, 11 (May 1974), 198-202.



## SITUATIONAL INFLUENCE IN INTERPERSONAL PERSUASION

Richard J. Lutz, University of California, Los Angeles  
Pradeep Kakkar, University of Pennsylvania

### Abstract

Situational influence on interpersonal persuasion processes is analyzed within the framework of decision process models of behavior, in particular the Fishbein behavioral intention model. Initially, relationships between the global situation and constructs in the Fishbein model are postulated, followed by an examination of current approaches to the study of situations. The interpersonal factor is identified as a key aspect of situations, and it is isolated as a key situational dimension for the study of interpersonal persuasion. A dyadic decision process model is then developed, with specific attention devoted to situational elements. Directions for research in three areas are suggested: situational taxonomy, situational influence on decision processes, and interpersonal persuasion as a dyad of decision makers.

In their excellent review of the personal selling literature, Davis and Silk (1972) report the results of a sales campaign initiated in Uruguay by a psychologist, Jacobo Varela, on behalf of an upholstery firm. As a part of the sales program, Varela attempted to duplicate, as closely as possible, key aspects of Asch's classic experiments on group pressure. Prospective retailers were invited to the company's offices in small groups to hear a sales presentation for ready-made curtains, a product which was contrary to some existing cultural norms. After a presentation utilizing a variety of visual aids (e.g., color slides, videotapes, actual product samples), the salesman immediately asked the prospective customer who appeared to be most favorable to the product (on the basis of facial expressions and other aspects of "body language") to express his opinion verbally, giving reasons for this opinion. Ultimately, the prospect was asked to commit himself to an order. The salesman then turned his attention to the prospect appearing next most favorable and repeated the process. Thus, each successive prospect faced a greater degree of group pressure to make an order. Similar to the results obtained by Asch, large proportions of initially resistant prospects were induced to make orders.<sup>1</sup>

The intent of this rather protracted introductory example is to underscore the importance of situational variables in the interpersonal influence process. Varela's program took the individual retailer from the comfortable confines of his own shop, where the salesman has a limited supply of persuasive resources, to the manufacturer's offices, complete with modern persuasive hardware and a carefully designed persuasive situation that used peer group pressure to enhance the receptivity of prospective customers. From this example, it is clear that, at least to some degree, the personal sales situation is a controllable variable, one that can be at least partially modified by the salesman to enhance his likelihood of success. However, before the situation can achieve full status as an explanatory variable for interpersonal persuasion and other forms of consumer behavior, methods and

<sup>1</sup>For a fuller description of Varela's program, see Zimbardo and Ebbeson (1969).

models must be developed to facilitate the study of situational influence.

The purpose of this paper is to propose a model for the assessment of situational influence on consumer decision making, particularly within the context of interpersonal persuasion. To accomplish this task, a conceptual model of situational influence on a single individual's decision process is presented, followed by a discussion of currently used approaches for the classification of situations. Finally, a model of interpersonal persuasion as a two-person decision process is developed, with specific reference to situational effects on this process.

### Situational Influence on the Consumer Decision Process

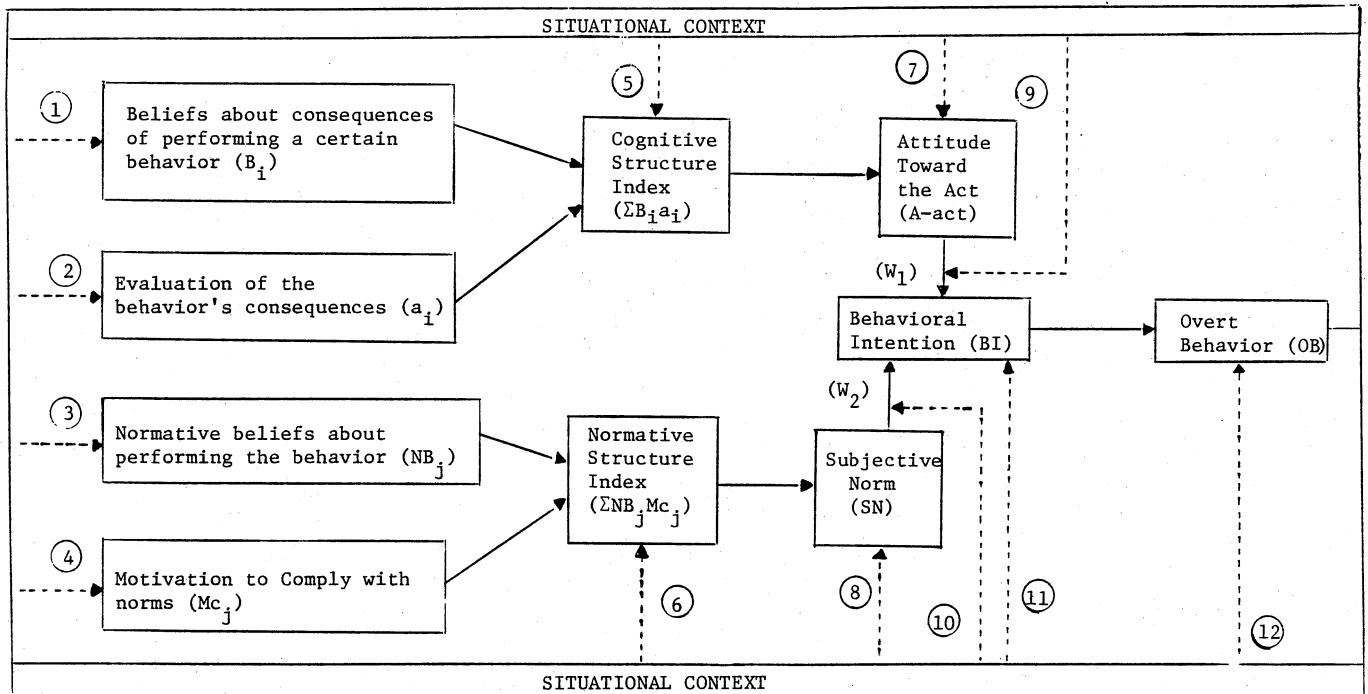
There are several models of the consumer decision process extant in the literature. For the purposes of the present discussion, Fishbein's (1957, 1975) modification of the theory of propositional control has been selected as representative of this general class of models. Three factors led to the selection of the Fishbein model: 1) more than any of its competitors, it has generated considerable interest and empirical testing by consumer researchers (See Ryan and Bonfield, 1975); 2) it makes some specific predictions regarding the locus of situational influence on parameters of the model; and 3) it is particularly suited to generating prescriptions for influencing the nature of the consumer decision process, thus making it a potentially valuable approach to the study of interpersonal persuasion.

In the interest of conserving space, a full presentation of the functional equations upon which the Fishbein theory rests will be omitted. The interested reader should see Ryan and Bonfield (1975) for a thorough presentation of the theory or, more conveniently, a paper by Lutz (1975) in this volume. Diagrammatically, the Fishbein theory appears as shown in Figure 1. Each box in the figure represents a theoretical construct; solid lines connecting the boxes are supposed causal linkages among the theory's constructs.<sup>2</sup> The dashed lines in the figure represent potential situational influences on constructs and relationships within the Fishbein theory. Thus the situation is seen as an exogenous variable which may exert influence on any of all of the endogenous variables shown in the figure. For the present, specific dimensions of the situation are not considered; rather, it is treated as a global construct. In a later section of this paper situational variables will be studied in more detail.

Each of the hypothesized relationships between the situation and Fishbein theory constructs has been numbered in the figure for the purposes of discussion. While space limitations prohibit a complete review of all

<sup>2</sup>Actually the flow of supposed causation shown in Figure 1 is only on configuration out of many possible; for instance, attitude (A-act) may exert causal influence "backward" on elements of cognitive structure. Nevertheless, the present representation is sufficient for the purposes of explicating situational influence.

FIGURE 1  
SITUATIONAL INFLUENCE ON THE CONSUMER DECISION PROCESS



literature pertinent to the hypothesized linkages, a brief rationale will be offered for each relationship. In some cases there are empirical findings to support the relationship. Table 1 provides a summary of illustrative research supporting the various relations and also gives brief hypothetical examples of each type of situational influence on the consumer decision process.

Turning first to situational influence on beliefs about consequences of behavior (#1 in the figure and table), the Asch (1956) conformity experiments provide evidence that group pressure can cause individuals to perceive stimuli differently than when the same stimuli are viewed in private. Similarly, Axelrod (1963) found that induced changes in subjects' moods led to substantial changes in their perceptions of products as being associated with those moods. These two studies represent only two situational factors possibly affecting  $B_i$ ; another possibility would be information available in the situation. In any case, there is little doubt that situational influence can be present in the  $B_i$  construct.

Situational influence on the  $a_i$  construct was also documented by Axelrod (1963). He found that the evaluations of nine different moods were significantly more favorable when subjects were actually in those moods than when they were not. In the same vein, Peak (1960) found need for achievement scores among students measured on the day of a quiz to be significantly higher than when the same group was measured on a day when no quiz was given. Both of these findings relate directly to the proposed situation influence on  $a_i$ .

In the same sense that the  $B_i$  construct is influenced by the presence of others, the  $NB_j$  construct (normative beliefs) should certainly exhibit the same tendency. Intuitively, it seems plausible that a person would be better able to estimate  $NB_j$  accurately when the significant others are physically present to tell him either verbally or through body language (e.g., frowning) what they expect of him. Again, the Asch (1956) study is used as justification for this hypothesized relationship,

as he demonstrated the pervasiveness of group influence on one type of beliefs. Presumably normative beliefs would be similarly affected by group pressures.

Motivation to comply with norms (#4 in the figure) can be seen as largely affected by the presence of others. This postulate is supported by the work of Kelman (1958), who pointed out that the degree to which individuals comply with group expectations is a function of the visibility of that conformity behavior of the group. Thus, for example, a person may be more motivated to comply with a salesman's expectations when the salesman is present.

Situational influence on the cognitive structure and normative structure indices (#5 and #6 in the figure) falls under the rubric of information processing research. Wright (1974) has shown that information processing "rules" used by consumers vary with time pressure and distractions in the situation. Thus the summative models shown in the figure may not apply equally well across all situations. Deviations from these models would appear in attenuated correlations between  $\Sigma B_i a_i$  and A-act between  $\Sigma NB_j Mc_j$  and SN.<sup>3</sup>

With respect to situational influence on A-act (#7 in the figure), Razran (1940) and Janis et al. (1965) demonstrated that a sort of halo effect sometimes operates in attitude change situations. If the situation is pleasant, attitude change in the direction advocated is enhanced, and vice versa if the situation is unpleasant. Thus a consumer can be expected to be more favorable toward products and also the influence of other people (#8 in the figure) in pleasant situations.

<sup>3</sup>The subjective norms construct, which is intended to be a summary measure of social influence, has only recently appeared formally as a part of the Fishbein model (Fishbein and Ajzen, 1975), although it had been alluded to informally for several years prior to that. Ryan (1975) has recently introduced a similar construct, Social Compliance, which also provides a summary measure of normative influence.

TABLE 1  
THEORETICAL SUPPORT FOR, AND EXAMPLES OF, SITUATIONAL INFLUENCE ON CONSUMER BEHAVIOR

Number of Relationship in Figure 1	Construct	Theoretical/ Empirical Support	Hypothetical Consumer Behavior Example
1	Behavioral Consequences ( $B_i$ )	Asch (1956) Axelrod (1963)	Consumer X sees shopping at Store A as "convenient" when he has an automobile available, but views Store A as "inconvenient" when he has no car available.
2	Evaluation of Consequences ( $a_i$ )	Axelrod (1963) Peak (1960)	Consumer X values economy highly when purchasing food for personal, private consumption, but not when purchasing food to serve to guests.
3	Normative Beliefs ( $NB_j$ )	Asch (1956)	Consumer X misperceives his family's expectations regarding which make of car to buy when they are not present; when they are present, he correctly perceives their expectations.
4	Motivation to Comply ( $Mc_j$ )	Kelman (1958)	Consumer X is quite anxious to go along with his family's expectations when they are present; when they are not, he does not care what they expect of him.
5	Cognitive Structure Index ( $\Sigma B_i a_i$ )	Wright (1974)	Consumer X combines cognitive elements in a linear, compensatory fashion when under no time pressure to reach a decision; when under time pressure, he reverts to simpler cognitive strategies.
6	Normative Structure Index ( $\Sigma NB_j Mc_j$ )		
7	Attitude (A-act)	Janis, Kaye and Kirschner (1965)	Consumer X, while consuming a delicious meal, is generally favorably predisposed to both products and other people; when hungry, however, his predispositions are not as favorable.
8	Subjective Norm (SN)	Razran (1940)	
9	Regression weight ( $W_1$ )	Fishbein (1967)	Consumer X's own attitude governs his behavioral intentions when he is alone; when he is with others, social pressure has a relatively larger influence on his intentions.
10	Regression weight ( $W_2$ )		
11	Behavioral Intentions (BI)	Sheth (1970) Triandis et al. (1972)	Consumer X definitely intends to purchase Brand Z when he goes shopping; however, he realizes that Brand Z is often unavailable, so he is only mildly certain that he will buy it.
12	Behavior	Brislin and Olmstead (1973) Wicker (1971)	Consumer X intends to buy Brand Z; however, when he looks for it in the store, it is out of stock, so he buys Brand Y instead.

Relationships #9 and #10 deal with the regression weights used in the Fishbein model to determine if behavioral intentions are under attitudinal ( $W_1$ ) or normative ( $W_2$ ) control. Fishbein (1967) has theorized that the relative magnitude of these two weights should shift depending upon the situation. For instance, in highly visible situations (e.g., personal selling),  $W_2$  might be expected to be relatively larger than under less visible conditions.

The situation may also exert direct influence on the formation of intentions (#11). For example, Sheth (1970) has posited that the "anticipated situation" is an important determinant of BI. This "anticipated situation" construct is essentially an assessment on the part of the individual as to what sorts of factors will ultimately be present when he attempts to perform the behavior in question. From another angle, Triandis et al. (1972) theorize that habit strength is a major determinant of BI. To the extent that a situation is highly familiar, the consumer may intend to engage in simple repetitive behavior with little cognitive activity.

Finally, the situation may exert influence directly on behavior. For example, Brislin and Olmstead (1973)

found that Wicker's (1971) "judged influence of extraneous events" construct controlled a large proportion of the variance in overt behavior, although slightly less than BI did. Nevertheless, there is little question that situational constraints frequently cause attenuation of the relationship between intentions and behavior.

From the preceding discussion, it can be seen that the Fishbein theory is highly situation-specific and thus requires situation-specific measurement procedures to be used effectively. Furthermore, not only does the situation have influence on the Fishbein theory at several points, but the resultant behavior is seen as having a reciprocating influence on the situation. Bandura (1974) has stated: "To the oft-repeated dictum, change contingencies and you change behavior, should be added the reciprocal side, change behavior and you change the contingencies" (Bandura, 1974, p. 866). This point takes on particular significance within the context of interpersonal persuasion, where the behavior of one person may substantially alter the situation within which the other person must act. This point will emerge more clearly in a subsequent section of this paper.

Up to this point, situational influence has been

discussed in a rather abstract manner, emphasizing the loci of such influence on the consumer decision process. However, to fully investigate situational influence, an adequate conceptualization and operationalization of the situation, as a construct, is necessary. The next section discusses some approaches to the problem of specifying the situation.

#### Specifying the Situation

Virtually every researcher to examine situational influence on behavior has pointed out the need for taxonomies of situations similar to the taxonomies of persons found in the literature on individual differences. Sells (1963b) summarizes the problem:

The most obvious need in evaluating the manifold encounter of organism and environment is a more satisfactory and systematic conceptualization of the environment. This implies a taxonomic, dimensional analysis of stimulus variables comparable to the trait systems that have been developed for individual difference variables . . .

In the absence of clear perception of the basic dimensions of the total stimulus situation, experimenters must have systematic information about relevant dimensions of the environment beyond the piecemeal, concrete, immediate variables customarily observed on the basis of experience. (p. 700)

In other words, it is not enough simply to describe in infinite detail a particular situation in which behavior is to occur. Rather, that situation should be characterized by locations along certain situation dimensions. This latter approach will allow the accumulation of knowledge about situational impact on behavior, while the former approach is likely to lead to a series of findings which are little more than anecdotal.

Several investigators, working in widely differing domains, have offered typological dimensions for the classification of situations. Table 2 summarizes some of these typologies. It is neither feasible nor desirable to discuss each of these typologies in detail at this time; nevertheless, there is one critical dimension along which all these typologies can be arrayed--that of degree of concreteness/abstractness. One of the most debated issues in situational research is whether the situation should be defined in "objective" or "subjective" terms (Belk, 1975; Lutz and Kakkar, 1975). Arguments on either side are compelling. On the one hand, Mausner (1963) has argued in favor of objectivity: "If one specifies the stimulus in terms of the nature of the receiver, lawfulness becomes impossible" (p. 107). But Rotter (1955), on the other hand, states: "The basic principle for classifying or categorizing a situation is psychological, that is, subjective . . ." (p. 259).

The typologies summarized in Table 2 represent virtually the entire spectrum ranging from objectivity/concreteness to subjectivity/abstractness. For instance, the dimensions employed by Rasmus (1970), Barker (1968), Porter and Lawler (1965), and Sells (1963a) are almost completely objective in nature, while those included by Rotter (1955), Murray (1938), Moos (1973), and Kakkar and Lutz (1975) are highly subjective. Sherif and Sherif (1956), Toffler (1970) and Belk (1974b) include some dimensions which are objectively defined and others which are more subjective in nature.

TABLE 2  
ALTERNATIVE SETS OF DIMENSIONS USED TO  
CLASSIFY SITUATIONS

Source	Domain	Situational Dimensions
Sherif and Sherif (1956)	Social Situations	Factors related to individuals involved, factors related to task or activity, factors related to the site, and interrelations among the three.
Toffler (1970)	General Situations	Place, people, physical surroundings, location in organizational network of society, context of ideas and information, span of time.
Mehrabian and Russell (1974)	General Situations	Pleasure, arousal, dominance (mediating emotional response variables).
Belk (1974)	Purchase	Physical surroundings, time frame, interpersonal surroundings, mood, goal direction.
Kasmar (1970)	Environmental Design	Meteorological variables, geographical variables, physical design variables.
Barker (1968)	Ecological Psychology	Descriptions of "behavior settings"--space, time, non-psychological objects.
Porter and Lawler (1965)	Organizational Behavior	Dimensions of organizational structure (e.g., size, centralization).
Sells (1963a)	General Situations	Personal characteristics of "milieu inhabitants."
Rotter (1955) Murray (1938)	General Situations	Reinforcement contingencies (e.g., affiliation, status).
Moos (1973)	Social Environments	Relationship dimensions, personal development dimensions, system maintenance and change dimensions.
Kakkar and Lutz (1975)	Consumption Situations	Social interaction factor, personal involvement factor, temporal commitment factor.

At the far end of the dimension in abstractness is the recent framework proposed by Mehrabian and Russell (1974) wherein situations are described solely by their impact on three mediating emotional response variables, thus circumventing the problem of defining the environment in either objective or subjective terms.<sup>4</sup> While this framework would probably not suffice as

<sup>4</sup>The interested reader is referred to Lutz and Kakkar (1975) for an initial application of this framework to consumer behavior.

the only means for conceptualizing situational influence, it can be a powerful tool for the explanation of situational influence when used in conjunction with other, more concrete typological schemes.

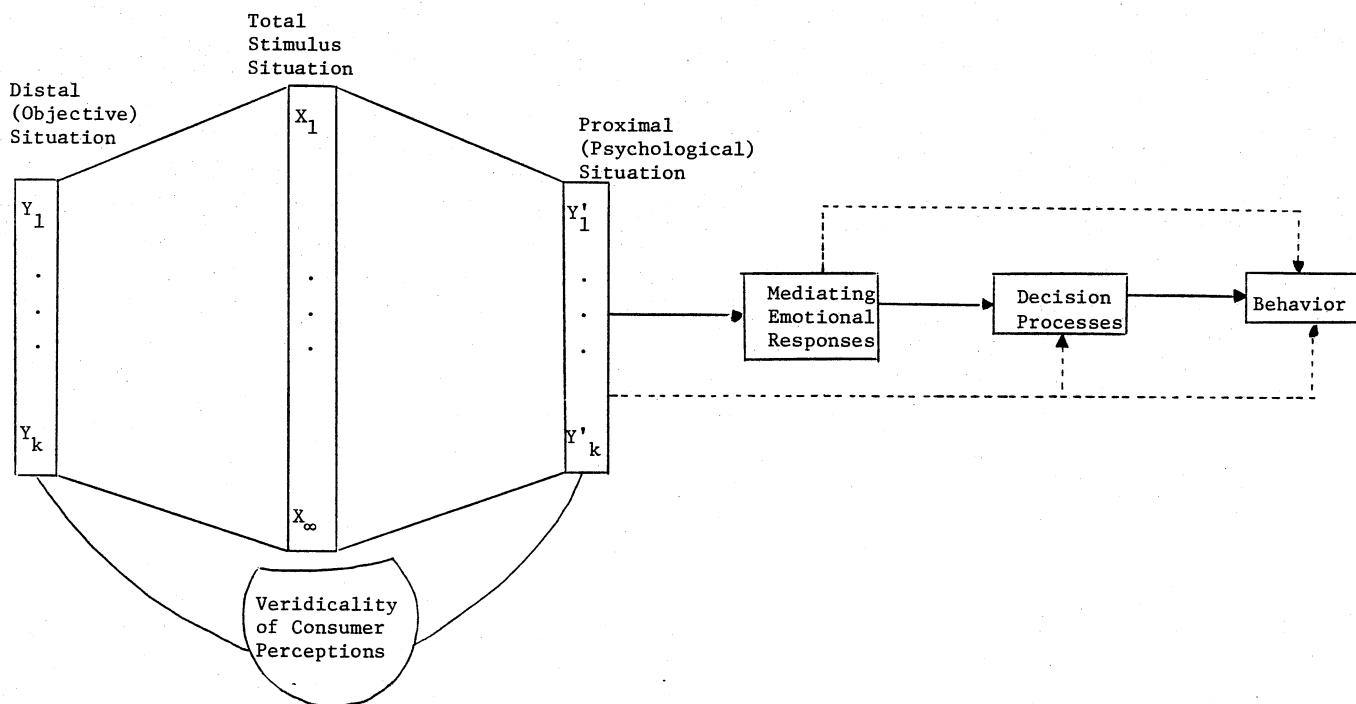
Several authors (e.g., Mausner, 1963; Belk, 1975) have advocated a unified approach in which objective and subjective definitions of the situation are combined to yield deeper insights into situational influence on behavior. By extending this notion to incorporate the emotional response mediators proposed by Mehrabian and Russell (1974), movement toward a "middle-range" theory (Ray, 1973) of situational influence on consumer decision processes is effected.

Figure 2 presents a model of situational influence which attempts to account for both the objective and subjective situation, as well as the mediating responses to the situation. First of all, the total stimulus situation is seen as consisting of virtually an infinite number of possible dimensions. To fully describe the total stimulus situation would be an onerous task at best and an impossible one in all likelihood. Nevertheless, it is the starting point for an analysis of situational influence.

interest here are the individual's perceptions of the situation rather than the objective characteristics per se. For instance, how many people does the person believe to be present, how large does the room seem, and how warm or cool does it feel to the person would be dimensions of the psychological situation. Viewing the two situation descriptions as vectors of situational variables, the veridicality of consumers' situational perceptions could be studied in a manner similar to that used by Brunswik (1952). That portion of Figure 2 is purposely patterned after his "lens model" to illustrate one interesting area for future situational research: the relationship between the objective and psychological situations. This point will be discussed in more detail later in the paper.

The psychological situation thus is seen as the translation of objective reality into psychological terms, which are regarded as internal stimuli to the individual, leading to the emotional responses which mediate the impact of the situation on decision processes and behavior. Mehrabian and Russell (1974) have presented an impressive array of evidence that individuals respond to situations along three primary dimensions of emotion: pleasure, arousal and dominance. Therefore,

FIGURE 2  
A MEDIATIONAL MODEL OF SITUATIONAL INFLUENCE



Fortunately, there is support for the delimiting of the situational definition to a subset of relevant dimensions. Belk (1974a) states: "... situation may then be defined as all those factors . . . which have a demonstrable and systematic effect on current behavior" (p. 156, emphasis added). Thus only those aspects or dimensions of the situation which are salient to the behavior under investigation are necessary for specification of the objective, or distal, situation, which as shown in the figure, can be described as  $k$  dimensions. These dimensions are objective descriptions of the situation, such as number of people present, size of room, temperature, etc.

In similar fashion, the psychological or proximal situation is described along the same  $k$  dimensions as is the objective situation. However, the data of

regardless of the specific dimensions of the psychological situation, levels of the three emotional response variables form the situational context within which decision processes and behavior occur, as shown previously in Figure 1.

The solid line connecting Mediating Emotional Responses and Decision Processes in Figure 2 is intended to show that the situation's impact on behavior is seen as predominantly filtered through decision process variables, as in Figure 1.<sup>5</sup> However, as was also indicated

<sup>5</sup>Kakkar (1975) has recently found significant relationships between pleasure, arousal and dominance and constructs within the Fishbein model shown in Figure 1.

in Figure 1, the situation may have direct impact on behavior. This possibility is shown in Figure 2 by the dashed line leading from Mediating Emotional Responses to Behavior. In like fashion, the dashed lines connecting Proximal Situation to Decision Processes and Behavior indicate that perceptions of the situation may directly influence those variables without being mediated by emotional reactions to the situation. These possibilities are discussed in more detail in the next section.

In sum, Figure 2 represents an overview of processes and mediating variables through which the situation is hypothesized to influence consumer behavior. Several research issues are raised by the form of the model in Figure 2--these will be discussed shortly.

### Situational Influence on Interpersonal Persuasion

In reviewing the situational typologies listed in Table 2, it is clear that a major element of the situation is the interpersonal dimension. The only two typologies which do not explicitly consider interpersonal influence are the highly abstract typology proposed by Mehrabian and Russell (1974) and the one used by Kasmar (1970) for the design of physical environments. Thus the interpersonal dimension would appear to be a very important one for the characterization of situations in general and especially for the study of interpersonal persuasion, where the salience of that dimension is undoubtedly heightened.

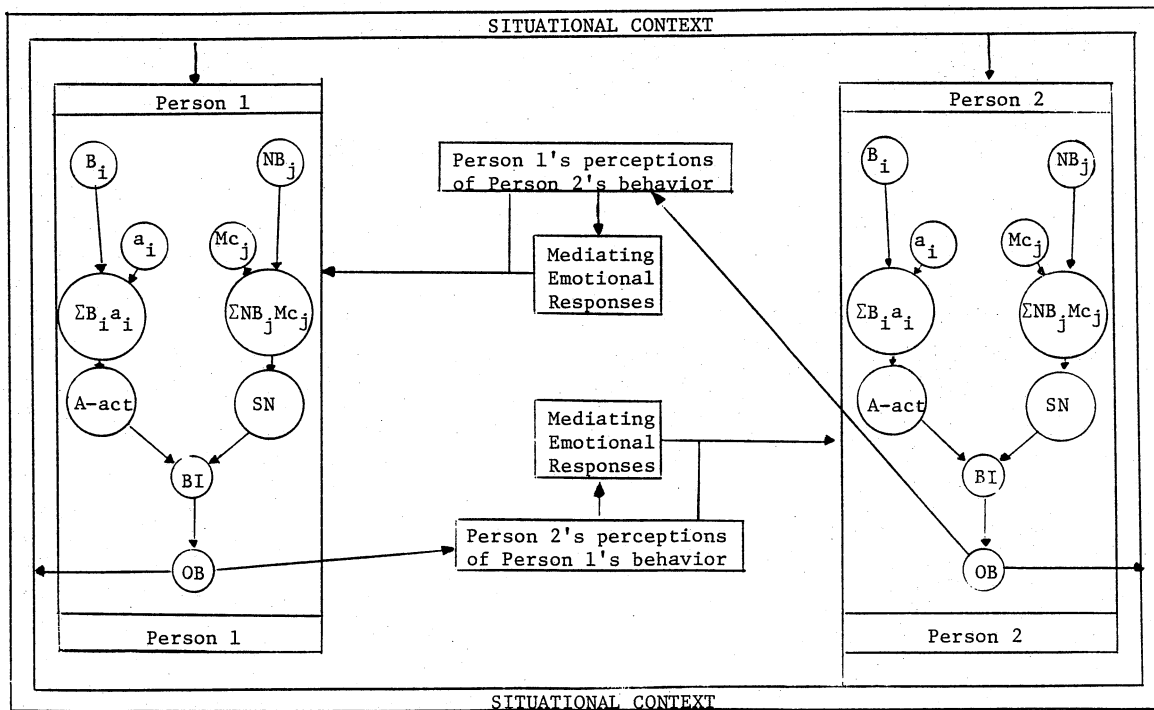
In order to provide a framework for the study of situational influence on interpersonal persuasion, an attempt has been made to combine key concepts of the decision process orientation shown in Figure 1 with the situational model of Figure 2. Using the persons in a dyadic interaction as the basic units of analysis, a model of situational influence on interpersonal persuasion is presented in Figure 3.

The model shown in Figure 3 is quite similar to that in Figure 1, with two important exceptions. First, the decision processes of two people, rather than just one, are included in Figure 3. This serves to illustrate the idea that in any interpersonal persuasion situation the two parties, i.e., the communicator and the receiver, are really decision makers. The communicator is making decisions as to what sort of appeals to use in order to influence the receiver to accept his position, product, etc. Thus, within the context of the Fishbein model, the behavior in question for the communicator is the use of some persuasive appeal. For the receiver, the behavior under analysis would be compliance with the influence attempt, e.g., purchase of a salesperson's product.

It will be noted that Person 1 and Person 2 in the figure are portrayed symmetrically. This is in acknowledgment of the fact that the roles of communicator and receiver are interchangeable throughout the persuasion process. For instance, the salesperson may initiate a sales appeal designed to influence the purchaser to buy a product at a certain price. The purchaser may then counter with a lower price, thus adopting the role of communicator in attempting to persuade the salesperson to accept the counter offer. This viewpoint is consistent with Thibaut and Kelley's (1959) theory of interaction outcomes, which employs dyadic interaction analysis to study behaviors such as two-person bargaining processes.

The second major difference between Figures 1 and 3 is the isolation of the interpersonal dimension of situational influence. Whereas in Figure 1 the entire situation was viewed as the context within which behavior occurs, in Figure 3 the behavior of the other person in the dyad is separated from the remaining dimensions of the situation. To avoid confusion, arrows indicating the various loci of situational influence on the decision process have been omitted in Figure 3. However, both Situational Context and the

FIGURE 3  
MICROFORMULATION OF THE INTERPERSONAL DIMENSION OF SITUATIONAL INFLUENCE



behavior of the other member of the dyad can serve as influences in any of the twelve relationships shown as dashed lines in Figure 1.

Furthermore, both Situational Context and dyadic behavior are subject to the same situational influence processes modeled in Figure 2. The psychological situation may act directly on individual decision processes and/or behavior or through the mediating emotional response variables.

A hypothetical example may serve to illuminate the use of the model in Figure 3 for analyzing interpersonal persuasion. Assume Person 1 is the communicator and is attempting to sell Person 2 a product. Person 1's first decision is whether or not to use a certain sales pitch. He holds beliefs ( $B_1$ ) about the use of this sales appeal (e.g., it works, it is unethical, it is difficult to perform) and evaluations ( $a_1$ ) of these attributes. Further, he holds some normative beliefs ( $NB_j$ ) with respect to the pitch (e.g., his boss expects him to use it, his children expect him not to), and he is more or less motivated to comply ( $Mc_j$ ) with these expectations. Each of the four constructs mentioned thus far is subject to situational influence, including influence of the behavior of the receiver. The communicator weighs all his beliefs ( $\Sigma B_1 a_1$  and  $\Sigma NB_j Mc_j$ ), arriving at an overall assessment of his feelings (A-act) and other's expectations (SN) regarding the use of the sales appeal. Depending upon the situation, either A-act or GSP carries more weight in the formation of his intentions (BI) to use the appeal. Finally, he makes a decision and uses the appeal (OB) to attempt to persuade Person 2.

Person 2 perceives Person 1's behavior along with other aspects of the situation. On the basis of this perception, which may or may not be an accurate reflection of Person 1's actual behavior (Figure 2), Person 1 may have an emotional response to the situation or may directly incorporate his perceptions into an inferential process which influences his beliefs, etc. Person 2's beliefs ( $B_2$ ) may include product attributes, both the number and level of which may be influenced by Person 1's behavior. Salient normative beliefs ( $NB_j$ ) are likely to include those of his family and the salesman himself. Many of Person 1's persuasive attempts may be directed at influencing attribute evaluations ( $a_1$ ) or motivation to comply ( $Mc_j$ ) with various referents. Similarly, Person 1 may influence the nature of Person 2's cognitive processes ( $\Sigma B_2 a_2$  and  $\Sigma NB_j Mc_j$ ), leading to impact on A-act and SN, the attitude and normative pressure toward purchase, respectively. Finally, Person 2 forms an intention (BI) to purchase or not and expresses this in overt behavior (OB), again subject to situational influence, both from Person 1 and the remainder of the Situational Context. If he decides not to purchase, then he immediately enters another decision process as the communicator, perhaps attempting to convince Person 1 to lower his price or to go away.

The reciprocal process outlined above will continue until one or the other of the two parties involved is either successful or leaves the situation. Use of the model shown in Figure 3 would enable the researcher to uncover which situational influences were instrumental in leading to the consummation of the interpersonal persuasion process, regardless of the nature of the final outcome or the number of iterations of the process required.

#### Future Research Directions

Future research on the impact of situational influence

on interpersonal persuasion falls into three main categories: 1) research on the nature of the situation and methods for specifying its dimensions; 2) research on situational influence on individual decision processes and behavior; and 3) research on the interpersonal persuasion dyad and the effects of the situation on it.

#### Research on the Situation

Referring back to Figure 2, several research questions emerge with respect to the situation per se. First of all, there is a need for taxonomical work within the consumer domain. Excellent starting points for this pursuit would be the work of Barker (1968) on behavior settings and Moos' (1973) analysis of treatment environments. The former approach is aimed at defining the objective situation, while the latter focuses on the psychological situation. Use of these two approaches in tandem to investigate situational dimensions in supermarkets, for example, may lead to interesting insights, particularly for public policy makers interested in introducing more product information into that situation.

A second area of needed research on the situation is the veridicality of consumer situation perceptions. Using Brunswik's (1952) lens model or the extension of it suggested by Dudycha and Naylor (1966) would be an appropriate methodological approach to this issue. The importance of this research cannot be overemphasized for the manager or policy maker hoping to influence consumer behavior by modifying situational dimensions. Unless the consumer perceives the situation along the same dimensions as the (objective) ones being manipulated, such strategies may prove fruitless.

The relationship between the psychological situation and measures of emotional response would provide useful policy inputs, provided that the emotional response variables relate to decision processes and behavior. For instance, it would be helpful to the marketer to know which aspects of the situation lead to high pleasure, low dominance and moderate arousal on the part of the consumer, for under these conditions Mehrabian and Russell (1974) have found the greatest tendency for subjects to approach the situation. By using this general finding to design store interiors, shopping centers, etc., the marketer may be able to increase the amount of time customers spend shopping in that environment. To do this, the relationship among emotional response and the psychological situation must be identified. Initial work in this direction has been reported by Kakkar and Lutz (1975).

#### Research on Situational Impact on Decision Processes

Figure 1 illustrates areas for research on the decision process. First, it should be determined which definition of the situation is most appropriate as the operationalization of "Situational Context." Of the three alternatives discussed in this paper--i.e., the objective situation, the psychological situation, and mediating emotional responses--the relationships of latter two to decision processes have recently been investigated by Kakkar (1975). More research is necessary to establish generalizations regarding the impact of situational variables on indices of cognitive structure, the formation of intentions, etc.

A second research thrust in this area would be the linkage of general patterns of findings relating situational dimensions to emotional response and general patterns of relationships between emotional response and decision processes. This linkage would again be a valuable one for managers and policy makers.

Research on the Interpersonal  
Persuasion Dyad

Many of the research issues in the two immediately preceding sections have a direct bearing on interpersonal persuasion as well: however, there are additional important questions raised in this context. First, the degree to which the interpersonal dimension of the situation is heightened in importance is a key question. If it is the dominant influence, then other features of the situation can be ignored by the marketer; if it is only one of several important influences, then the marketer will have to be concerned with other controllable dimensions of the situation. Initial work in this area has been reported by Wilson et al. (1972).

Secondly, the combination of a situational perspective and a decision process orientation may provide a fruitful area for extensions of social exchange theory, particularly Thibaut and Kelley's (1959) theory of interaction outcomes. Expansion of the model shown in Figure 3 to incorporate multiple behavioral options would be a first step in this direction.

In conclusion the model proposed in this paper for the study of situational influence on interpersonal persuasion process can be viewed as an intersection of three major classes of behavior theories: situational influence, decision process models, and social exchange theory. Through this combination perhaps all three theories can gain in richness. The benefits to those engaged in interpersonal persuasion would include an increased understanding of the processes through which it works and prescriptions for the improvement of the processes through which it works and prescriptions for the improvement of the effectiveness of interpersonal persuasive attempts.

References

- Asch, Solomon E., "Studies of Independence and Conformity: A Minority of One Against a Unanimous Majority," Psychological Monographs, 70, no. 9 (1956).
- Axelrod, Joel N., "Induced Moods and Attitudes Toward Products," Journal of Advertising Research, 3 (June 1963), 19-24.
- Bandura, Albert, "Behavior Theory and Models of Man," American Psychologist, 29 (December 1974), 859-69.
- Barker, Roger G., Ecological Psychology: Concepts and Methods for Studying the Environment of Human Behavior (Palo Alto: Stanford University Press, 1968).
- Belk, Russell W., "An Exploratory Assessment of Situational Effects in Buyer Behavior," Journal of Marketing Research, 11 (May 1974a), 156-63.
- \_\_\_\_\_, "Situational Influence in Consumer Behavior." Unpublished manuscript, Temple University (1974b).
- \_\_\_\_\_, "The Objective Situation as a Determinant of Consumer Behavior," in Advances in Consumer Research, Vol. 2, ed. M. J. Schlinger (Chicago: Association for Consumer Research, 1975), pp. 427-37.
- Brislin, Richard W., and Kim H. Olmstead, "Examination of Two Models Designed to Predict Behavior from Attitude and Other Verbal Measures," Proceedings Eighty-first Annual Convention, American Psychological Association (1973), pp. 259-60.
- Brunswick, E., "The Conceptual Framework of Psychology," in International Encyclopedia of Unified Science, ed. R. Carnap and C. Morris (Chicago: University of Chicago Press, 1952).
- Davis, Henry L., and Alvin J. Silk, "Interaction and Influence Processes in Personal Selling," Sloan Management Review, 13 (Winter 1972), 59-76.
- Dudycha, Linda Weathers, and James C. Naylow, "Characteristics of the Human Inference Process in Complex Choice Behavior Situations," Organizational Behavior and Human Performance, 1 (1966), 110-28.
- Fishbein, Martin, "Attitude and the Prediction of Behavior," in Readings in Attitude Theory and Measurement, ed. M. Fishbein (New York: John Wiley, 1967), pp. 477-92.
- \_\_\_\_\_, and Icek Ajzen, Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research (Reading, Mass.: Addison-Wesley, 1975).
- Janis, Irving L., Donald Kaye and Paul Kirschner, "Facilitating Effects of 'Eating-While-Reading' on Responsiveness to Persuasive Communications," Journal of Personality and Social Psychology, 1 (1965), 181-86.
- Kakkar, Pradeep, "Situational Influence on Emotional Response, Cognitive Structure and Cognitive Process," Unpublished Doctoral Dissertation, University of California, Los Angeles, 1975.
- \_\_\_\_\_, and Richard J. Lutz, "Toward a Taxonomy of Consumption Situations," Combined Proceedings, American Marketing Association, 1975, forthcoming.
- Kasmar, J. V., "The Development of a Usable Lexicon of Environmental Descriptors," Environment and Behavior, 2 (1970), 153-69.
- Kelman, Herbert C. "Compliance, Identification, and Internalization: Three Processes of Attitude Change," Journal of Conflict Resolution, 2 (1958), 51-60.
- Lutz, Richard J., "Conceptual and Operational Issues in the Extended Fishbein Model," in Advances in Consumer Research: Vol. 3, ed. B. B. Anderson (Cincinnati: Association for Consumer Research, 1975, forthcoming).
- \_\_\_\_\_, and Pradeep Kakkar, "The Psychological Situation as a Determinant of Consumer Behavior," in Advances in Consumer Research, Vol. 2, ed. M. J. Schlinger (Chicago: Association for Consumer Research, 1975, pp. 439-53).
- Mausner, Bernard, "The Specification of the Stimulus in a Social Situation," in Stimulus Determinants of Behavior, ed. S. B. Sells (New York: Ronald Press, 1963), pp. 107-16.
- Mehrabian, A., and J. A. Russell, An Approach to Environmental Psychology (Cambridge: MIT Press, 1974).
- Moos, Rudolph H., "Conceptualizations of Human Environments," American Psychologist, 28 (August 1973), 652-65.
- Murray, Henry, Explorations in Personality (New York: Oxford University Press, 1938).
- Peak, Helen, "The Effect of Aroused Motivation on Attitudes," Journal of Abnormal and Social Psychology, 61 (1960), 463-68.
- Porter, L., and E. Lawler, "Properties of Organization Structure in Relation to Job Attitudes and Job



- Behavior," Psychological Bulletin, 64 (1965), 23-51.
- Ray, Michael L., "Psychological Theories and Interpretations of Learning," in Consumer Behavior, Theoretical Sources, ed. S. Ward and T. S. Robertson (Englewood Cliffs, N.J.: Prentice-Hall, 1973), pp. 45-117.
- Razran, G. H. S., "Conditioned Response Changes in Rating and Appraising Sociopolitical Slogans," Psychological Bulletin, 37 (1940), 481.
- Rotter, Julian B., "The Role of the Psychological Situation in Determining the Direction of Human Behavior," in Nebraska Symposium on Motivation, ed. M. R. Jones (Lincoln: University of Nebraska Press, 1955), pp. 245-69.
- Ryan, Michael J., "A Test of Associations in a Causal Chain Derived from Fishbein's Behavioral Intention Model," Unpublished working paper, University of Alabama, 1975.
- Ryan, Michael J., and E. H. Bonfield, "The Fishbein Extended Model and Consumer Behavior," Journal of Consumer Research, 2 (1975), forthcoming.
- Sells, S. B., "Dimensions of Stimulus Situations Which Account for Behavior Variance," in Stimulus Determinants of Behavior, ed. S. B. Sells (New York: Ronald Press, 1963a), pp. 3-15.
- \_\_\_\_\_, "An Interactionist Looks at the Environment," American Psychologist, 28 (September 1963b), 696-702.
- Sherif, Muzafer, and Carolyn W. Sherif, "Varieties of Social Stimulus Situations," in Stimulus Determinants of Behavior, ed. S. B. Sells (New York: Ronald Press, 1963), pp. 82-106.
- Sheth, Jagdish N., "An Investigation of the Relationships Among Evaluative Beliefs, Affect, Behavioral Intention and Behavior," Unpublished working paper, University of Illinois at Urbana-Champaign, 1970.
- Thibaut, J. W., and H. H. Kelley, The Social Psychology of Groups (New York: John Wiley, 1969).
- Toffler, Alvin, Future Shock (New York: Random House, 1970).
- Triandis, Harry C. et al., The Analysis of Subjective Culture (New York: John Wiley, 1972).
- Wicker, Allan W., "An Examination of the 'Other Variables' Explanation of Attitude-Behavior Inconsistency," Journal of Personality and Social Psychology, 19 (1971), 18-30.
- Wilson, David T., H. Lee Mathews, and John F. Monky, "Attitude as a Predictor of Behavior in a Buyer-Seller Bargaining Situation: An Experimental Approach," Combined Proceedings, American Marketing Association, Chicago, 1972, pp. 390-95.
- Wright, Peter, "The Harassed Decision Maker: Time Pressure, Distraction and the Use of Evidence," Journal of Applied Psychology, 1975, forthcoming.
- Zimbardo, Philip, and Ebbe B. Ebbesen, Influencing Attitudes and Changing Behavior (Reading, Mass.: Addison-Wesley, 1969).

CONSUMER DECISION MAKING IN NATURALISTIC SETTINGS:  
SALESMAN-PROSPECT INTERACTION

Richard W. Olshavsky, Indiana University

The potentially great influence personal selling can have upon consumer decision making has been recognized for a very long time by marketers. A great deal of time, money and effort is expended in recruiting, training, evaluating and compensating salesmen. Much has been written about salesmen-prospect interactions. But most of these efforts have taken the form of prescriptive formulas on "How To Become A Successful Salesman." Very few systematic studies have been performed; and most of these have been concerned with the influence of preselected variables such as age, sex or personality on the outcome of the transaction. Only a couple studies have attempted to examine the complex dyadic interaction at the process level. Consequently, in spite of the importance of the topic, today we know very little about consumer decision making in naturalistic settings (see Capon, Holbrook and Hulbert [1975] for a comprehensive and recent review).

There are at least two reasons for the neglect of research on this important area of consumer behavior. The first reason concerns the enormous complexity of the behavior of interest. Two (or more) people interacting over long periods of time (sometimes over multiple encounters) about products or services which are usually technically complex can generate a stream of behavior which is extremely complex. It is a challenging task to develop a methodology which permits reliable and valid measures of the important aspects of this behavior and to analyze the resulting data in meaningful ways.

The second important reason for the neglect of this important area of consumer behavior pertains to the lack of a theory of salesman-prospect decision making. The absence of such a theory makes it difficult to identify the important determinants of the dyadic interaction process and how these determinants interact over time to produce the behavior of interest.

These two important sources of difficulty are actually very closely interrelated. Good theory leads to the identification of relevant determinants of behavior which in turn facilitates the identification of appropriate methodology. Good methodology, on the other hand, produces results which facilitate the development of better theory.

#### Purpose of Paper

The purpose of this paper is to advocate, in the spirit of a workshop, the general outline of a theory of salesman-prospect interaction and to illustrate the type of methodology that is implied by this new theory.

#### A Theory of Salesman-Prospect Interaction

To explain the behavior of a salesman-prospect dyad, it is posited that it is first necessary to know the decision process of each member of the dyad with respect to the task at hand. This implies that a theory of individual decision making exists which is capable of generating behavior in considerable detail. One theory which meets this requirement is the Newell and Simon information processing theory (IPT) of human

cognition (Newell and Simon, 1972; Olshavsky, 1975).

Newell and Simon's IPT accounts for human cognition in terms of an interaction which occurs among a person's goals, internal representation of the task environment and strategy. It is assumed that sequences of observed behavior are generated by the execution of the selected strategy. The strategy itself is nothing more than a particular configuration of elementary information processes which the person is assumed to be capable of interpreting and executing. IPT further postulates that cognitive man is to be viewed as one instance of a general class of information processing systems. The essential features of such a system are: receptors and effectors which transmit symbolic information between the system and the external environment, a central processor which can interpret and execute a limited set of elementary information processes and a memory which is capable of storing symbols and symbol structures over long periods of time.

These basic postulates are turned into testable hypotheses by the application of the theory to a particular problem solver, concept learner or decision maker in a particular task environment. For the particular instance, the problem solver's goals, internal representation of the objective task environment and his strategy are specified in terms of a precise data structure and a precise organization of elementary information processes to be executed. These processing rules and internal representation of the task environment are then expressed in a formal computer language and the logical consequences are derived by performing the indicated operations according to the specified rules. The resulting computer program, therefore, becomes a microtheory of the behavior of interest; at the same time, it is a detailed and testable model of that theory.

Although IPT has been developed and applied primarily to task environments involving single individuals, it is possible to extend this theory to situations where two or more individuals are interacting in the same external environment. Clarkson (1968) for example demonstrated how this theory could be extended to small group decision making by positing that the behavior of the group is the result of the two or more information processing systems acting interdependently over time. In group decision making, two or more individuals are required to work together such that the end result of the joint deliberation is a single decision; each participant's behavior becomes part of the task environment for the other participants. Interpersonal influence in this setting is represented as the effect one or more members of the group have upon the goals, strategies or internal representations of the others. The extent of such influence was taken by Clarkson as a measure of the degree of leadership expressed by each participant. Clarkson was able to show that a valid model of group decision making could be developed by first developing models of each participant and by identifying the type and extent of interpersonal influence which occurred over multiple group decisions.

Dyadic interaction between salesman and consumer differs in some important respects from small group decision making. Participants in a group decision making situation all share essentially the same basic goal. Moreover, they are likely to have similar internal representations of the task environment and similar strategies for attempting to achieve this common goal. In the case of salesman-prospect interactions, however, the goals of the participants will rarely be the same. The salesman's goal will generally be to sell the prospect some product or service from his existing inventory while the consumer's goal will be less definite and more idiosyncratic.

Likewise, while both salesman and prospect interact in the same objective environment, their task environments differ considerably. The salesman's task environment consists of his particular inventory of goods or services along with a particular prospect; the consumer's task environment consists of the salesman and the salesman's inventory of goods or services. In general we may expect the salesman to have a more accurate and complete internal representation of his task environment than the consumer does.

We may also expect major differences in the type of strategies elicited by each participant in an effort to achieve his goals. Whereas the salesman may have a well organized, pre-planned strategy for accomplishing a sale, the consumer's strategy may be poorly organized and in some cases nonexistent. In my study of appliance shoppers, e.g., it was found that several consumers simply turned the choice task over to the salesman (Olshavsky, 1973).

Just as in group decision making, interpersonal influence in salesman-prospect interactions can be represented in terms of the influence each participant has upon the other's goals, internal representations and strategies of each participant and the interpersonal influence which occurs between participants over time in the form of a computer model. Through simulation, the model can then be tested for validity. Validated models can then be used as the basic data from which generalizations about consumer decision making in naturalistic settings can be made. As an additional bonus, generalizations will also be forthcoming about salesman strategies and their effectiveness with different types of consumers. (Valuable insights about other marketing management considerations such as the appropriate depth and breath of inventories may also be gained from such models.)

#### Illustrative Methodology

The research objectives prescribed by IPT are to identify the goals, internal representations and strategies of human problem solvers. In order to expose these hidden mental processes as well as possible, Newell and Simon and others who have adopted the IPT have resorted to a technique called protocol analysis. Protocol analysis involves having the subject under investigation talk aloud as he attempts to perform the specified cognitive task. Close examination of the protocol data leads, by inference, to specific hypotheses about the goals, internal representations and strategies of individual subjects.

A group problem solving situation creates some serious methodological problems for an IPT researcher. Because of the potential impact that the exposed thought process of one participant might have upon the thought process of others in the group, protocol analysis is not a viable methodology unless it is used creatively. Clarkson, e.g., had individual participants talk aloud

in a pre-task which was similar to the task they solved as members of a group. Another methodological problem of group or dyadic behavior arises from the very interactive nature of the process itself. This interactive process may result in important and rapid changes in the participants' goals, internal representations and strategies such that continuous monitoring of each participant is required. A final methodological problem that arises in the case of salesman-prospect interaction in naturalistic settings is that a great deal of relevant information can be directly "read" from the external environment by visual inspection. For instance, an appliance shopper can directly observe the styling, color, size and prices of alternatives and a salesman can observe the approximate age, social class and income of the shopper. All of these considerations pose serious methodological problems for the application of protocol analysis to salesman-prospect interactions.

In the interest of stimulating research of the type prescribed by IPT, a specific methodology will now be advanced to illustrate how some of these methodological difficulties might be overcome. The specific methodology to be proposed is a variation on the Willett and Pennington (1966) field study. In the Willett and Pennington study, audio recordings were made of actual transactions between salesman and prospect as they took place on the retail floor for several types of appliances. These audio recordings were later supplemented by post purchase interview data. The Willett and Pennington procedure has certain definite advantages. By dealing with actual transactions in actual settings, findings from such studies are likely to have high external validity. Moreover, since the subjects in this study were unaware that they were being recorded, they were not influenced by the experimental procedures.

From the point of view of an IPT theorist, however, the Willett and Pennington procedure has certain serious shortcomings. In particular, the audio recordings do not reveal the thought processes as well as protocol data can. The lack of detail in the audio recordings does not permit easy inference concerning the participant's strategy or their internal representation. And it is difficult with audio data alone to detect changes in the important aspects of cognitive behavior and to detect the factors which produce these changes. Audio recordings also fail to reveal the type and amount of information potentially available to participants by direct visual inspection.

The first modification that should be made to the Willett and Pennington procedure therefore is to supplement the audio recordings with video recordings. The second modification is to have each participant separately provide a detailed analysis of the transaction immediately after the transaction is over. This retrospective analysis should use the audio-video recording the way today's sports announcers and athletes use the instant replay for diagnosis of sports events. Such an analysis could provide the necessary additional detail concerning each participant's goals, internal representations and strategies throughout the transaction. The use of the audio-video recording will stimulate detailed analysis and will minimize any possible memory loss or distortion. Each participant's detailed remarks should be recorded for later analysis by the researcher.

These two types of data, the original audio-visual recordings and the audio recordings of the detailed retrospective analysis by the participants, can then be subjected to analysis by the researcher. Original goals, internal representations and strategies can be inferred as well as any changes which occur in them

and the determinants of these changes. Hypotheses can then be formalized in the form of a single comprehensive model of both the salesman and the consumer and the interactions which occurred between them. The resulting model can then be tested by making detailed comparisons between the model's output and the actual transactional data. Validated models of several different salesmen and prospects can then serve as data points from which generalizations about consumer and salesman decision making in naturalistic environments can be made.

In order to facilitate the success of this proposed procedure, it is suggested that initial studies be limited in the following ways: only one type of product or service, or very few, should be investigated, only dyads should be studied as opposed to groups of three or more, realistic settings should be used but unnecessary distractions and interruptions should be minimized, e.g., by using an actual retail floor after closing hours and one shot transactions should be investigated as opposed to multiple exposure transactions. After some experience has been gained with the proposed methodology and theory, it will be an easy matter to relax these simplifying restrictions.

#### Closing Remarks

In closing, I would like to emphasize that a salesman can and often does exert a profound influence upon consumer decision making for many products and services. Until we come to grips with the complexities of dyadic and small group transactions, we will never have a truly complete understanding of consumer decision making and our ability to make meaningful suggestions for marketing management will be greatly limited. It is my hope that the theoretical and methodological proposals outlined in this paper, brief as they are, will stimulate further empirical research and theoretical development in this important area of consumer behavior.

#### References

- Capon, Noel, Morris Holbrook and James Hulbert. The Selling Process: A Review of Research. (Mimeograph, submitted for publication to the Journal of Marketing, 1975.)
- Clarkson, Geoffrey P.E. Decision Making in Small Groups: A Simulation Study. Behavioral Science, Vol. 13, 1968, 288-305.
- Newell, Allen and Herbert A. Simon. Human Problem Solving. Englewood Cliffs, N.J.: Prentice-Hall, 1972.
- Olshavsky, Richard W. Customer-Salesman Interaction in Appliance Retailing. Journal of Marketing Research, Vol. X (May 1973), 208-12.
- Olshavsky, Richard W. Implications of an Information Processing Theory of Consumer Behavior. Proceedings, American Marketing Association, Fall conference (1975), forthcoming.
- Willett, Ronald P. and Allen Pennington. Customer and Salesman: The Anatomy of Choice and Influence in a Retailing Setting. Proceedings, American Marketing Association, 1966, 598-616.

## BUYER-SELLER INTERACTION: A CONCEPTUAL FRAMEWORK

Jagdish N. Sheth, University of Illinois

### Abstract

This paper attempts to provide a comprehensive conceptualization of the buyer-seller interaction process. The basic postulate under the conceptualization is that the quality of interaction is a function of the compatibility between the buyer and the seller with respect to both the style and the content of communication. After defining the dimensionalities of style and content, a number of personal, organizational and product-specific factors are described as determinants of style and content of communication in buyer-seller interaction process.

### Introduction

A review of the literature in the area of buyer-seller interaction process points out at least three dimensions of the state of the art (See Evans, 1963; Davis and Silk, 1972; Hulbert and Capon, 1972; O'Shaughnessy, 1972; and Webster, 1968 for summaries and reviews of the knowledge in the area).

First, the extent of empirical research on the buyer-seller interaction process is relatively sparse suggesting considerably less interest in this area at least among the academic researchers. While there is considerable talk about the mysteries of the super-salesman and some good research in the area of selection and training of sales representatives in industrial marketing, the vital linkage of the buyer-seller interaction process remains yet to be systematically researched.

Second, whatever empirical research one finds in the area is highly sporadic and ad hoc. Most of it consists of attempts to extend specific hypotheses borrowed from the behavioral sciences to describe and explain process of buyer-seller interaction. These consist of several similarity hypotheses related to the backgrounds and physical characteristics of the buyer and the seller, and the reliance on the Yale School of thought on personal communication including impact of source, message and channel factors (Howard and Sheth, 1969; Capon, Holbrook and Hulbert, 1975). Consequently, the area of buyer-seller interaction is replete with numerous hypotheses, interesting observations and considerable degree of contradictory or unrelated research findings.

Third, there is a conspicuous absence of any comprehensive conceptualization or theory of buyer-seller interaction. It seems no one has as yet attempted to go beyond reviewing the literature in order to sort out existing evidence and to reconcile inconsistent or contradictory findings by offering a comprehensive or holistic perspective to the problem area.

A comprehensive perspective of the buyer-seller interaction process seems timely and can serve several useful functions. It will encourage more systematic and realistic research which takes into account many interdependent phenomena relevant to understanding the buyer-seller interaction process; it will probably point out new areas of research by providing insights which can only come from a comprehensive perspective; finally it is likely to discourage research in what

may prove to be irrelevant or less useful subareas. Often, research in a growing area tends to localize in a very narrow issue losing sight of the many other unexplored and more useful aspects within it. Witness the recent experience in the area of attitude structure and specifically the controversy about the judgmental rules a person utilizes in processing multiattribute information.

Accordingly, the purpose of this paper is to attempt a comprehensive conceptualization of the buyer-seller interaction process. It is hoped that such a conceptualization will generate additional insights into the problem area and encourage more selective and concerted research.

### Overview

The conceptual framework suggested in this paper is comprehensive and abstract enough to include buyer-seller interactions in both household and organizational marketing. In other words, it is capable of explaining the process of buyer-seller interaction which takes at the retail outlets for consumer goods as well as between sales representatives and purchasing agents of formal organizations.

It is also comprehensive enough to include all types of buyer-seller interactions. These can be interpersonal (face to face), written or even telecommunication in nature. It is surprising to note how written and telecommunication buyer-seller interactions have been ignored in past research activities.

The conceptual framework developed in this paper has consciously avoided extending any particular well-known theory of interpersonal communication from the behavioral sciences. Often, such blind extensions have proved less useful in the past (Sheth, 1974b). Instead, attempt is made to conceptualize the area from a managerial perspective and selectively choose as many theories and hypotheses from behavioral sciences as seem relevant to provide insights into why and how some buyer-seller interactions work to the satisfaction of both the parties and others don't.

The basic postulate underlying the conceptual framework summarized in Figure 1 is that whether a specific buyer-seller interaction will or will not work is a function of two distinct dimensions of interaction. The first dimension is the content of communication representing the substantive aspects of the purposes for which the two parties have got together. It entails suggesting, offering, promoting or negotiating a set of product-specific utilities and their expectations.

While the dimensions of product-specific utilities will be described in detail later in the paper, it is sufficient to note here that often the expectations offered by the seller and desired by the buyer for a specific product or service do not match resulting in failure of the interaction transaction to be consummated successfully and satisfactorily.

A second dimension of buyer-seller interaction determination is the style of communication. It represents the format, ritual or mannerism which the buyer and the

seller adopt in their interaction. The style of interaction reflects the highly individualistic preferences and normative expectations of the buyer and the seller about the process of interaction itself. Much of the search for the supersalesman is often localized in identifying the style of interaction of highly successful salesman in organizational marketing.

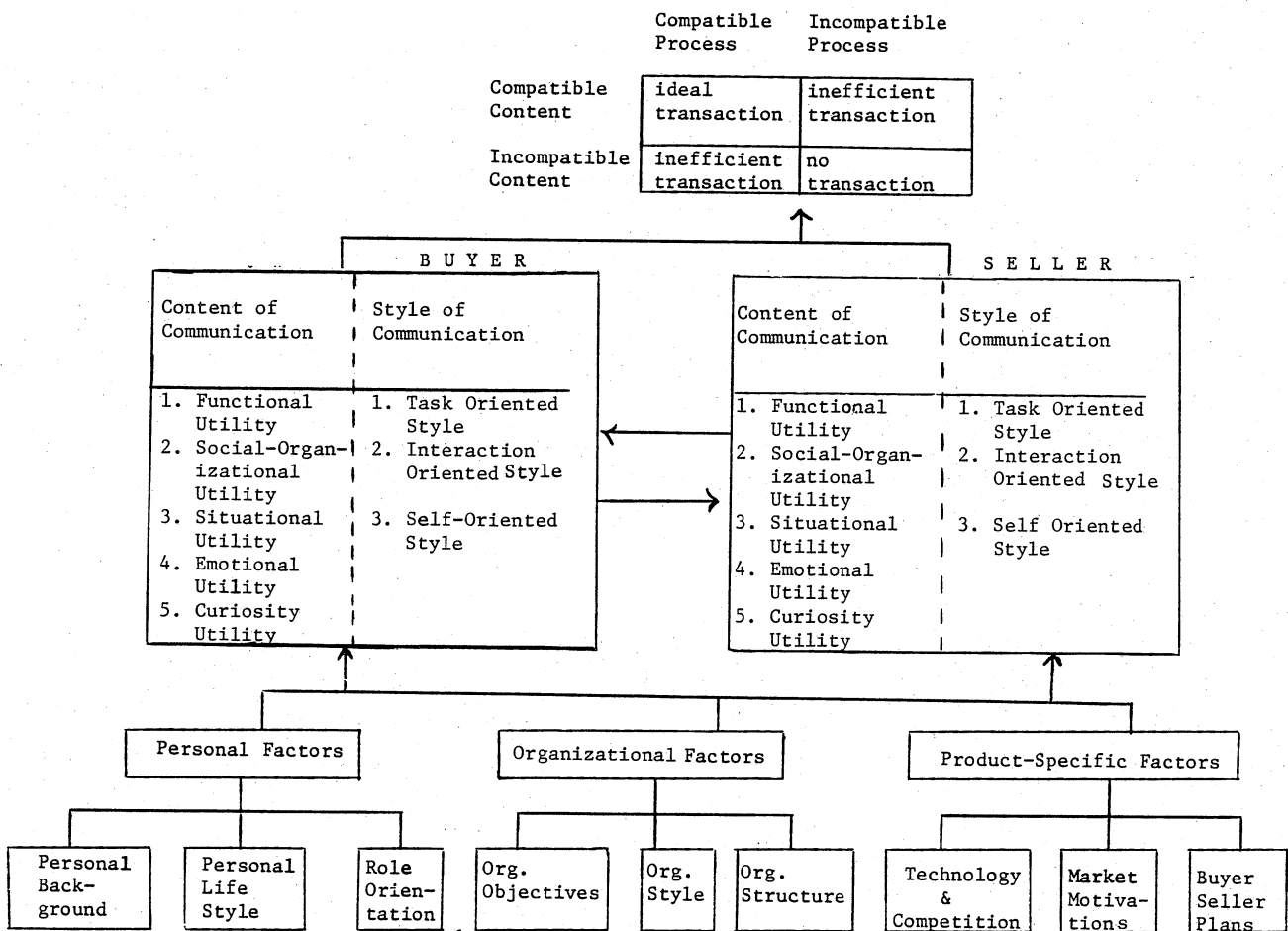
The buyer-seller interaction process itself is treated as a transaction which can have multiple effects or consequences. Comparable to the impact of advertising (Sheth, 1974a), the buyer-seller interaction is presumed to perform any of the following five functions: (a) increase awareness of each other's expectations about the product or service; (b) remind each other's past satisfactory transactions and their behavioral outcomes; (c) reinforce each other's behavior related to the sale of the product or service; (d) precipitate behavioral actions on each other's part by intensifying expectations; and (e) persuade each other to change their respective expectations.

Whatever the objective, a satisfactory interaction transaction between the buyer and the seller will occur if and only if they are compatible with respect to both the content and style of communication. In all other situations, the interaction transaction is

presumed to be less than ideal. In figure 1, a two by two classification of interaction transaction is provided as a very simple framework to understand the impact of incompatibility with respect to style and content of communication. For example, if the buyer and the seller are compatible with respect to style but not with respect to content of communication, it is argued that while a dialogue will continue between the two parties, the actual sale may not be consummated due to difference in product expectations. Either the interaction process will be terminated or negotiations will take place to change each other's product expectations. On the other hand, if the buyer and the seller are compatible with respect to content but not the style of communication, it is argued that either the process will be terminated or even if the sale is consummated there will be negative feelings about each other's style or manner of interaction resulting in an unsatisfactory transaction. Finally, when both the style and the content are incompatible between the buyer and the seller, not only will there be no transaction culminating in a sale, but there are likely to be negative side effects of complaints, bad word of mouth about each other, and distrust of each other.

Both the style and content of buyer-seller communication are determined by a number of personal, organizational

FIGURE 1. A CONCEPTUAL FRAMEWORK OF BUYER SELLER INTERACTION



and product-related factors. For example, the personal life styles and backgrounds will often determine the style of communication the buyer or the seller chooses to engage in. Similarly, organizational training and orientation will also mould the buyer or the seller with respect to the style of communication he is expected to engage in. Finally, the content of communication is likely to be determined by product-related variables such as market motivations, buyer and seller plans and technology or competitive structure of industry.

#### Content of Communication

While it is obvious that any incompatibility with respect to what the buyer wants and what the seller offers in a product or service will be detrimental to consummating a sale, it is more interesting and useful to identify dimensions and sources of content incompatibility. Based on a recent model of individual choice behavior (Sheth, 1975), it is proposed that underlying buyer-seller expectations about a product or service, there lies a five dimensional utility space. The five dimensions represent different types of product-related utilities which the buyer desires and the seller offers to each other. Each type of utility is briefly described below:

1. Functional Utility. It represents product's utility which is strictly limited to its performance and which defines the purpose of its existence and classification as a type of good or service. For example, the functional utility associated with an instant breakfast can be described in terms of taste, convenience, nutrition and calories. Similarly, the functional utility associated with a passenger car tire can be defined in terms of mileage, blow out protection, traction, handling and ride. The functional utility is often measured in terms of a person's expectations on a number of product-anchored attributes or evaluative criteria. It is presumed to be a complex function of positive and negative expectations on multiattribute profiles. In this paper, we treat functional utility as one dimension of product utility and ignore for a moment the question of its own dimensionality.

2. Social-Organizational Utility. Sometimes a product or service acquires social-organizational connotations or imageries independent of its performance or functional utility. This is due to its consistent identification with a selective set of socioeconomic, demographic or organizational types. Such identification with a selective cross-section of household or organizational buyers tends to impute certain utilities or disutilities in the product or service producing an imagery or a stereotype. For example, cigarettes are often consumed due to their social imagery even though they may be functionally harmful. Certain products are, therefore, used for their prestige and not so much their performance. The existence of social-organizational utility in a product or service is also prevalent in organizational buyer behavior especially with respect to those products and services which are directly associated with the organization man. This is not surprising in view of the fact that there exists an organizational stratification of people working in organizations comparable to social stratification of households based on organization structure, hierarchy, and power distribution.

3. Situational Utility. It represents a product's utility which is derived from existence of a set of situations or circumstances. The product or service has no intrinsic or independent utility and will

not be offered or bought without the presence of circumstances which create its need. The situational utility is often strong among those products or services which are consumed on an ad hoc basis rather than on a continuous basis. For example, the utilization of the services of the priest for marriage ceremony or the lawyer for divorce proceedings tend to be nonrepetitive by and large. Similarly, a housewife may buy a product or service as a gift item due to a very specific situation or occasion such as graduation or marriage. Organizations often tend to use the services of professionals on an ad hoc basis because of a specific project. Many of the capital expenditure items and highly specialized professional skills have greater degree of situational utility in them. It is extremely important to identify situations and activities which add to the utility of the product or service.

4. Emotional Utility. Sometimes a product or service evokes strong emotive feelings such as respect, anger, fear, love, hate or aesthetics due to its association with some other objects, events, individuals or organizations. The strong emotive feelings are therefore generalized to the product or service resulting in a different type of utility or disutility. For example, some Jewish buyers tend to refrain from buying German products because of strong emotional feelings they arouse as reminders of the German Nazi movement. Similarly, many Hindus refrain from eating beef due to strong emotive feelings anchored in religious tenets. While one would expect less prevalence of emotive utility in organizational products or services than in household products or services, this is not borne out by empirical research. Organizations also tend to manifest emotive behavior as is evidenced in international trade and cross-national negotiations.

5. Curiosity Utility. The fifth type of utility often present in both household and organizational products or services is related to novelty, curiosity and exploratory needs among individuals. Based on the assumption that man constantly seeks out new, different things due to either satiation with existing behavior or due to boredom inherent in highly repetitive tasks, certain new products or services acquire additional utilities which are not intrinsic to their performance. These products or services are both offered and sought largely due to their novelty and to satisfy a person's curiosity arousal. They have a very short life cycle and often degenerate as fads or fashions.

Each product or service has a vector of the five types of utilities described above. Furthermore, both the buyer and the seller will have certain expectations about the product or service on these five types of utilities. It is not at all uncommon both in household and organizational marketing to learn that the specific utility expectations of the buyer and the seller do not match resulting in some form of incompatibility with respect to content of interaction.

The degree of incompatibility can be measured by performing a dimensional analysis of the vectors of buyer-seller expectations. For example, we can locate the vectors of buyer and seller expectations in a five dimensional space, and measure the degree of incompatibility as a function of the distance between the buyer and the seller points located in the space. The greater the distance between the buyer and the seller points in space, the greater the incompatibility with respect to the content of communication. Presuming the equivalence between Euclidian distance and psychological incompatibility, the degree of incompatibility can be measured as follows:

$$D_{BS} = \sqrt{\sum_{j=1}^5 (b_{B_j} - b_{S_j})^2} \quad (1)$$

where  $D_{BS}$  = Distance or incompatibility between Buyer and Seller

$b_{B_j}$  = Buyer's expectations with respect to jth type of utility

and  $b_{S_j}$  = Seller's expectations with respect to jth type of utility

The distance between the buyer and the seller will determine to what extent they are matched with respect to content of communication. Since the buyer in a free enterprise system has the economic buying power, it is presumed that the seller will often adapt or change his offerings in such a way as to minimize the distance. However, it is often not true in reality because the seller also attempts to change the location of buyer expectations in the space by persuasive communication strategies or sales tactics.

Who will make the adjustment is clearly a function of who has the greater power in the buyer-seller relationship. While the buyer has the economic power, the seller often has greater technical expertise to offset buyer's power. As a very broad generalization, it is likely that in a buyer's market, the seller is more likely to change in the long run. In the seller's market, it is more likely that the buyer will change or adapt. In all other cases, tactics of persuasion, negotiations and bargaining are likely to emerge as consequences of buyer-seller interaction.

#### Style of Communication

The vast literature on group dynamics and interpersonal relationships in small groups (Bass, 1960; Heider, 1958; Homans, 1961), provides an excellent source to discuss the concept of style of interaction. As mentioned before, it refers to the format, ritual and mannerism involved in buyer-seller interaction. While we will rely heavily on research in group dynamics, it is important to keep in mind that the dimensionalities of style of interaction discussed here are common to nonpersonal interactions such as via telecommunication or postal systems. The style of interaction is presumed to be three dimensional. The specific dimensions are described below:

1. Task-Oriented Style. This style of interaction is highly goal oriented and purposeful. The individual is most interested in the efficiency with which the task at hand can be performed so as to minimize cost effort and time. Any activity during the interaction process which is either not task-oriented or inefficient is less tolerated by the individual who prefers the task-oriented style. The buyer or the seller who prefers this style of interaction often tends to be mechanistic in his approach to other people.

2. Interaction-Oriented Style. The buyer or the seller who prefers this style of interaction believes in personalizing and socializing as an essential part of the interaction process. In fact, preference for this style of interaction is often manifested at the loss or ignoring of the task at hand. The buyer or the seller motivated by the interaction-oriented style is often compulsive in first establishing a personal relationship with the other person and then only getting involved in the specific content of interaction.

3. Self-Oriented Style. This style reflects a person's preoccupation with himself in an interaction situation. He is more concerned about his own welfare

and tends to have less empathy for the other person. He is often unable to take the other person's perspective and views all aspects of interaction from his own selfish point of view. The concepts of self-preservation, self-survival and self-emulation tend to dominate this style of interaction.

It is also not uncommon to find situations in which the buyer and the seller are incompatible with respect to style of interaction. Given a three-dimensional vector of style of interaction, it is possible to measure the extent of incompatibility with the following Euclidian distance:

$$D_{BS} = \sqrt{\sum_{j=1}^3 (C_{B_j} - C_{S_j})^2} \quad (2)$$

where  $D_{BS}$  = Distance between Buyer and Seller on style of interaction

$C_{B_j}$  = Buyer's orientation with respect to jth type of style of interaction

$C_{S_j}$  = Seller's orientation with respect to jth type of style of interaction.

The greater the distance between the buyer and the seller points in the style space, the more incompatible they will be with respect to style of interaction.

Unlike content of interaction, it is more difficult to change or adapt with respect to style of interaction. This is largely because the style orientations of individuals are often deep rooted in personality variables, early socialization processes and personal life styles. It is, therefore, difficult to discuss who should make changes in what situation in the buyer-seller interaction process. If the style of interaction is highly incompatible between the buyer and the seller, it is probably best to terminate interaction and attempt to link the right types of sellers with the buyers in the interaction process.

#### Determinant Factors

Both the style and the content of buyer-seller interaction are determined by a set of exogeneous factors. These are classified into three categories: (a) personal factors anchored to the individuals involved in the interaction; (b) organizational factors anchored to the respective organizations the buyer and the seller belong to. Even in household marketing, we believe there are organizational factors not only associated with the seller but also with the buyer in so far as a typical household has some organizational structure, no matter how implicit it may be; (c) product-related factors anchored to market motivations, competitive structure and buyer-seller plans. We will briefly describe some of the more salient variables in each category. However, it is beyond the scope of this paper to treat them exhaustively or even attempt to specify their causal influences on the style and content of interaction.

1. Personal Factors. The personal factors are likely to determine the style of interaction each individual prefers. Among many personal factors, there seems to be some consensus among the researchers with regard to the following specific variables. The first one is the demographic, socioeconomic and organizational background of the individual. These include physical characteristics such as sex, race, height, weight etc. as well as both generalized education and special skills acquired by the individual. A second specific variable is the individual's life style. It reflects the moulding of the individual over time as a function of socialization and personality development. The third specific variable is the role orientation of the individual with



respect to the interaction process. It includes expectations and performance of specific roles on the part of the salesman such as a consultant, order taker, informer, persuader, etc.

2. Organizational Factors. Organizational factors often determine both the style and the content of interaction. The organization often recruits, selects, trains and prepares the buyer or the seller with respect to both the content and style of communication. The organizational factors which account for variability among organizations in their degree of controlling the content and style of interaction are organization objectives, organization style and organization structure. The content will be heavily influenced by organization objectives and to some extent by organization structure. Similarly, each organization has explicit or implicit style of management often dictated by the top man in the organization. The organization style is likely to influence the personal style of communication of the seller or the buyer.

3. Product-Specific Factors. The product-specific factors are more likely to determine the content rather than the style of interaction. While there are many specific factors one can include in the list, we will isolate three specific factors which seem more relevant and interesting. The first factor, of course, relates to market motivations. It refers to the generalized needs, wants and desires customers have for which the specific product is more or less relevant. The second factor relates to buyer and seller plans. The buyer has certain plans in his mind about the specific use he is likely to make use of the product. Similarly, the seller has certain plans with respect to market differentiation and customer segmentation. The product expectations of the buyer and the seller are likely to be heavily determined by their respective plans. The third factor is anchored to the supply side of the product. It refers to the technological and competitive leadership the seller has in that product category. The product expectations and utilities especially in regard to functional, situational and curiosity utilities are more likely to be determined by technology and competition prevalent in the industry.

The three types of determinants of style and content of interaction are extremely relevant to isolate individual differences among buyers and sellers, product differences for the same buyer or seller, and organizational differences for the same product. They essentially serve the function of reducing all the buyer-seller interactions to a common base by partialling out the effects of personal, organizational and product differences.

#### Conclusion

This paper has attempted a comprehensive conceptualization of the buyer-seller interaction process based on the presumption that whether or not there will be a satisfactory interaction will depend on whether the buyer's and the seller's style as well as content of interaction match. To the extent they do not match, the interaction is likely to be either terminated or will entail negative side effects.

Knowledge of mismatch between the buyer and the seller either with respect to style or with respect to content will require managerial corrective actions. These actions may take the form of modifying sales appeals, retraining salespeople, reassignment of salesmen as well as changes in recruiting and selection of personnel.

#### References

- Bernard M. Bass, Leadership, Psychology and Organizational Behavior (New York: Harper & Brothers, 1960).
- Noel Capon, Morris Holbrook and James Hulbert, "The Selling Process: A Review of Research" (unpublished paper, 1975).
- Franklin B. Evans, "Selling as a Dyadic Relationship- A New Approach," The American Behavioral Scientist, 6 (May, 1963), 76-79.
- Harry L. Davis and Alvin J. Silk, "Interaction and Influence Process in Personal Selling," Sloan Management Review, 13 (Winter, 1972), 54-56.
- Fritz Heider, The Psychology of Interpersonal Relations (New York: John Wiley & Sons, 1958).
- George Homans, Social Behavior: It's Elementary Forms (New York: Hartcourt, Brace & World, 1961).
- John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior (New York: John Wiley & Sons, 1969).
- James Hulbert and Noel Capon, "Interpersonal Communication in Marketing: An Overview," Journal of Marketing Research, 9(February, 1972), 27-34.
- John O'Shaughnessy, "Selling as an Interpersonal Influence Process," Journal of Retailing, 47(Winter, 1971-72), 32-46.
- J. N. Sheth, "Measurement of Advertising Effectiveness: Some Theoretical Considerations," Journal of Advertising, 3(January, 1974), 6-11.
- J. N. Sheth, "The Future of Buyer Behavior Theory," in J. N. Sheth (ed.) Models of Buyer Behavior (New York: Harper & Row, 1974).
- J. N. Sheth, "Toward a Model of Individual Choice Behavior," paper presented at the ESOMAR Seminar on Modeling, June 11, 1975.
- Frederick E. Webster, Jr., "Interpersonal Communication and Salesman Effectiveness," Journal of Marketing, 32(July, 1968), 7-13.

SELF-PERCEPTION AS A MEANS OF PERSONAL INFLUENCE:  
THE FOOT-IN-THE-DOOR TECHNIQUE

Brian Sternthal, Northwestern University  
Carol A. Scott, Ohio State University  
Ruby Roy Dholakia, Northwestern University

Abstract

Two experiments were conducted regarding the foot-in-the-door phenomenon as a personal influence strategy. The first experiment examined the efficacy of the foot-in-the-door under different levels of incentive. The second tested the adequacy of self-perception theory as an explanation for the foot-in-the-door effect.

The preponderance of the research pertaining to interpersonal influence has focused on the efficacy of various mass communication strategies. In part, the relative neglect of individual influence contexts reflects the lack of appropriate experimental paradigms to study the phenomenon. In part too, the neglect is attributable to the belief that strategies demonstrated to be effective in mass communication settings will also be appropriate in personal contact situations. However, in light of the unique attributes of personal contact, such as the availability of immediate feedback from message recipients, it may well be that certain influence strategies which fail in mass communication contexts will be successful in individual influence situations.

This paper examines the effectiveness of one individual influence strategy, the "foot-in-the-door" technique. The foot-in-the-door, as it is commonly employed in personal selling, entails asking individuals to make a relatively small purchase (or trial) in the hope that compliance with the request will increase the likelihood of compliance with subsequent larger purchase solicitations. In contrast to mass persuasion strategies, which are predicated on the assumption that behavior can be modified by influencing its cognitive precursors, the foot-in-the-door technique involves direct modification of behavior. More specifically, it is based on the self-perception proposition which states that individuals come to know their attitudes by inferring them from their own behavior and the circumstances in which that behavior occurs. In turn, the attitudinal dispositions acquired in this manner guide subsequent action (Bem, 1972).

Several field experimental studies have demonstrated the success of the foot-in-the-door technique. In their seminal investigation, Freedman and Fraser (1966) observed that individuals who agreed to comply with a small request evinced a greater willingness to comply with a subsequent larger request than people asked to comply with only the larger request (i.e., cold called). Moreover, this foot-in-the-door effect emerged whether or not the small request involved the same issue and task as the large request.

Subsequent investigations have generally confirmed the foot-in-the-door finding, and have specified the conditions under which it was likely to be observed. Pliner, Hart, Kohl, and Saabi (1974) reported that compliance with a large request, donating money to charity, was greater when it was preceded by a small or moderate request than by no request. However, the dollar amount given by those contributing did not vary according to whether or not experimental participants had complied

with a previous small request. Baron (1973) replicated the foot-in-the-door finding, but only when males served as experimenters and the initial request was very small. Cialdini (in review), however, failed to replicate the foot-in-the-door phenomenon, probably because small and large requests were made in close temporal proximity.

These foot-in-the-door findings are generally congenial to self-perception theory (Bem, 1972). Indeed, Freedman and Fraser's interpretation of their own data was in attribution terms:

...Once he has agreed to a request, his attitude may change. He may become in his own eyes, the kind of person who does this sort of thing, who agrees to requests made by strangers, and who takes action on things he believes in, who cooperates with good causes (Freedman and Fraser, 1966, p. 201).

In effect, it is hypothesized that the foot-in-the-door is effective because people use their own behavior as a cue regarding their attitudinal dispositions. Since external pressure for the initial behavior is assumed to be minimal, people infer a positive attitude from their compliant behavior, which in turn guides subsequent action.

In a test of this explanation, Snyder and Cunningham (1975) found that, in accord with the self-perception prediction, people who complied with an initial small request (answering 8 questions in a telephone survey) were more willing to comply with a subsequent large request (answering 30 questions) than subjects who were initially asked to comply with a very large request (answering 50 questions), or subjects in the no initial request condition. Similarly, Baron (1973) found that non-compliance with an initial large request led to reduced compliance with a subsequent smaller request, whereas compliance with an initial small request led to greater compliance with the subsequent request. No differences were observed between the initial large request group and no first request group.

Thus, individuals' actions are causally related to their prior behavior. Further, it has been suggested that individuals incorporate the circumstances in which behavior occurs into the attribution rules they employ in performing causal analysis of their own behavior. Specifically, for behavior to result in belief inferences it must be self-attributed, or perceived to be elicited by the person's disposition toward the stimulus. Circumstantial forces can attenuate the probability of self-attribution of behavior and the resulting belief inferences and behavior. If the behavior is perceived to be elicited by plausible external causal factors present in the situation (e.g., coupon, incentive, high credibility source), the individual should discount internal motivations as the cause of his behavior and no belief inferences should be made (discounting principle, Kelley, 1971). In a field experiment, Uranowitz (1975) tested the self-perception explanation under two different conditions of external justification. Consistent with the self-perception prediction, women asked to watch the experimenter's packages when justification was low (experimenter had to retrieve a dollar) exhibited greater willingness to help

a second experimenter than when justification for watching his packages was high (retrieve his wallet) or when there was no initial contact.

In sum, the literature demonstrates that the foot-in-the-door technique is an effective means of influence relative to cold calling. Furthermore, these data are congenial to self-perception theory; the performance of a behavior increases the likelihood of more substantial behaviors of the same genre, providing that individuals do not perceive circumstantial factors to have caused initial compliance. However, none of the studies reviewed have examined the foot-in-the-door technique in a commercial setting where incentives are likely to be necessary to induce consumers to engage in even small behaviors. Nor has there been a direct test of the self-perception hypothesis that past and future behaviors are mediated by attitudinal dispositions.

The present research involves two experiments. In Experiment I, the effectiveness of the foot-in-the-door technique is examined in a typical marketing setting, and the extent to which self-perception theory can explain these findings is assessed. Experiment II provides a more rigorous test of the self-perception explanation.

#### Experiment I: Foot-in-the-Door in a Commercial Context

##### Overview

In Experiment I, participants were given one of four incentives to enhance compliance with a small request. Two weeks later all participants were recontacted and asked to comply with a large request. A control group was asked to comply only with the large request. In accord with self-perception theory, it was hypothesized that individuals contacted for an initial small request (trial) would be more likely to comply with a subsequent large request than those contacted only for the large request.

It was also expected that participants asked to comply with an initial small request without benefit of an incentive would be more likely to comply with a second large request than those participants offered an incentive for initial trial. Although not performed in the context of the foot-in-the-door paradigm, studies pertaining to incentive strategies support this hypothesis. Doob and his coworkers (1969) introduced new brands of inexpensive, frequently-purchased consumer products in experimental stores at a discounted price and in control stores at the regular (competitive) price. After a nine-day period, discounts were removed and the effects on product sales observed. As expected, experimental stores outsold control stores during the discount period. When the discount was eliminated, however, this advantage was lost. In fact, by the end of a four-week observation period the control stores surpassed experimental stores in sales. Similarly, Aaker (1969, 1972), used regression analysis of panel data to investigate the effects of an initial purchase under deal conditions on subsequent brand acceptance. His findings suggest that purchase on deal may be negatively associated with acceptance, although the results across product categories were less than univocal.

Evidence for the hypothesis that behavior predicated on incentives undermines the performance of subsequent behavior also obtains from studies employing an over-justification paradigm (e.g., Calder and Staw, 1975; Deci, 1971; Kurlanski, et al., 1975; Lepper and Greene, 1975). Overjustification is operationalized as behavior which is both intrinsically attractive and externally rewarded. The consistent finding of these studies is

that the provision of a reward or incentive for performing an enjoyable behavior results in a decrease in intrinsic interest in subsequently engaging in that behavior.

##### Method

A weekly community newspaper serving the northwest suburbs of Chicago was selected for study. The paper had just been introduced to the market and was suitable for the experiment, since it could be tried on a limited basis and since the cost of trial was sufficiently low to insure a reasonable rate of compliance with an initial request.

Four hundred and twenty subjects were selected from a list of potential subscribers provided by the publisher and were randomly assigned to one of five groups: four experimental groups and one control group. Subjects in the experimental treatment conditions were called by the newspaper's regular telephone solicitation staff and asked to accept a two-week trial subscription to the paper at either the regular price (50¢), at half price (25¢), for free, or for free plus receipt of a premium gift (a coupon worth 50¢ at a well-known fast food restaurant). Two weeks later, all participants (all experimental participants, regardless of whether they accepted the trial offer, and control group participants) were called and asked to subscribe to the newspaper on a regular basis. Subjects were required to select at least a six-months subscription, which cost \$4.

The independent variable was the amount of incentive offered for trial subscription acceptance (none, 25¢, 50¢, 50¢ plus a premium gift). The percent of subjects in each treatment who agreed to subscribe on a regular basis served as the dependent variable. The control group, contacted only for the regular subscription request, was included to provide a measure of the response that would be obtained with no trial offer of any kind.

##### Results

According to self-perception theory, experimental participants who were given a two-week trial offer should evince greater rates of subscription than those merely asked to subscribe. This hypothesis was not supported by the data. When all trial offer conditions are combined, and compared with the no trial group, the differences in subscription behavior were nonsignificant ( $\chi^2=1.26$ , d.f.=1,  $p>.15$ )<sup>1</sup> Thus, trial per se does not enhance the likelihood of repeat purchase behavior.

When responses are categorized by the amount of incentive offered for trial, the effects of trial are somewhat different. As Table 1 indicates, the fifty percent discount treatment produced a significantly greater rate of subscription than the no-trial control treatment ( $\chi^2=3.0$ , d.f.=1,  $p>.05$ )<sup>2</sup>. The lowest rate of sub-

<sup>1</sup>All significance levels reported for 2x2 contingency tables are based on one-tailed tests, since the comparisons were performed to test directional hypotheses.

<sup>2</sup>Despite the fact that paired comparisons were planned rather than conducted in a post hoc fashion, it might be contended that inflated error rates due to the non-orthogonality of the contrasts caused intertreatment differences to be significant. To guard against this possibility an analog of Scheffe's procedure for multiple comparisons was used (Maraciullo, 1966). However, since this procedure did not change the interpretation of the findings it is not reported here.

scription occurred for the free trial group. No significant differences were found between the regular price and fifty percent discount treatments ( $X^2=.73$ , d.f.=1,  $p>.20$ ). Contrary to expectations, the free plus premium gift condition produced almost as great a subscription rate as the fifty percent discount treatment, although it was not significantly greater than the no-trial control treatment ( $X^2=1.06$ , d.f.=1,  $p>.15$ ).

TABLE 1  
Subscription Behavior of Each Trial Offer and Control Groups (Total Sample)

	<u>% Not Subscribing</u>	<u>% Subscribing</u>	
Regular Price trial	86.8% (105)	13.2% (16)	(n=121)
50% Discount trial	81.9% (95)	18.1% <sup>ab</sup> (21)	(n=116)
Free trial	96.1% (49)	3.9% (2)	(n=51)
Free & Premium trial	83.3% (35)	16.7% <sup>c</sup> (7)	(n=42)
No Trial - Control	91.0% (91)	9.0% (9)	(n=100)
Total	87.2% (375)	12.8% (55)	N=430

Overall  $X^2 = 8.41$ , d.f.=3,  $p=.08$

- <sup>a</sup>Significantly greater than control group,  $p<.05$   
<sup>b</sup>Significantly greater than free trial group,  $p<.02$ .  
<sup>c</sup>Significantly greater than free trial group,  $p<.05$ .

The above analyses were conducted using all participants, irrespective of their trial acceptance behavior. This was necessary to avoid the self-selection bias that would have been present had only those who complied with the trial request been included in the analysis. From a theoretical perspective, however, there is good reason to examine the responses of only the trial compliers. Specifically, Lepper, *et al.* (1975) state that overjustification effects may be evident only when there is some initial interest in the activity. Where no initial interest exists, the smallest amount of incentive necessary to elicit behavior should be used. In addition, since discounts are usually offered to encourage product trial by those persons who would not otherwise do so, it would be useful to know how successful this strategy is in influencing repeat purchases.

The subscription behavior of the trial compliers in each incentive condition is shown in Table 2. Here, the superiority of the fifty percent discount treatment is clearly demonstrated; the rate of compliance obtained in this condition is significantly greater than that obtained from any larger incentive treatment, though not significantly greater than the regular price condition rate ( $X^2=.56$ , d.f.=1,  $p>.25$ ). Again, the lowest subscription rate occurred under the free trial condition. Further, only the regular and fifty percent discount groups produced significantly greater subscription rates than the no-trial control group ( $X^2=9.14$  and  $18.96$ , respectively, d.f.=1,  $p<.01$ ).

#### Discussion

The results of this study have several practical implications. First, the finding that the fifty percent

TABLE 2  
% Trial Acceptors Complying With Subscription Request

	<u>% Subscribing</u>	
Regular Price trial	32.4% <sup>ab</sup> (11)	(n=34)
50% Discount trial	44.1% <sup>abc</sup> (15)	(n=34)
Free trial	2.8% (1)	(n=36)
Free & Premium trial	18.8% <sup>d</sup> (6)	(n=32)
No Trial - Control	9.0% (9)	(n=100)
Total	17.8% (42)	N = 236

Overall  $X^2$  (excluding control group) = 18.08, d.f.=3  $p<.01$ .

- <sup>a</sup>Significantly greater than free trial group,  $p<.01$ .  
<sup>b</sup>Significantly greater than no-trial control,  $p<.01$ .  
<sup>c</sup>Significantly greater than premium trial group,  $p<.05$ .  
<sup>d</sup>Significantly greater than free trial group,  $p<.05$ .

discount was the only trial offer that was significantly more effective than the cold calling approach suggests that some small incentive is necessary to gain the greatest benefits from a foot-in-the-door strategy. Second, although the fifty percent trial and the premium trial (i.e., free trial plus 50¢ gift certificate) were about equally effective in generating subscriptions, the latter offer was four times as expensive (\$1 versus 25¢) and thus is less preferred. Third, the data suggest that the most efficient procedure in implementing the foot-in-the-door technique entails confining call-backs to those people who accepted the trial offer; merely recontacting people had no effect on subscription behavior. Indeed, those persons who rejected the trial offer were no more likely to subscribe to the newspaper than the control group which was not made a trial offer (Table 3). It would probably be more efficient to use a promotional mix that combined advertising and a small incentive to arouse product interest and trial acceptance.

TABLE 3  
Subscription Behavior of Trial Offer Non-Acceptors

	<u>% Not Subscribing</u>	<u>% Subscribing</u>	
Trial Offer - Non-Acceptors	93.3% (181)	6.7% (13)	n=194
No-Trial Offer - Control	91.0% (91)	9.0% (9)	n=100
Total	92.5% (272)	7.5% (22)	N=294

$X^2=.24$ , d.f.=1,  $p>.80$ .

The results of this study provide only equivocal support for self-perception theory. As expected, the free trial offer yielded less compliance with the subscription request than did smaller incentives. Furthermore, the failure to observe significant differences in compliance behavior between the regular and half-price treatment groups, though not predicted by self-perception theory, does not constitute damaging evidence. Unlike the potent

discounting cues used in previous research (e.g., drugs or a sizable reward), the incentive used in the present research may have been of a sufficiently small magnitude so that it served as a positive reinforcer rather than a discounting one.

More problematic from a self-perception standpoint is the finding that the premium offer was as successful in inducing subscriptions as the fifty percent discount. Several factors may account for this finding. The small difference between the two groups when all subjects were included is due primarily to the large number of non-compliers with the fifty percent discount trial request relative to the premium trial offer (70.7% versus 23.8%). When only those who complied with the trial request are examined, the fifty percent discount resulted in more subscriptions than did the premium offer. These data imply that despite similar rates of compliance with the subscription request in the fifty percent discount and premium treatments, the reason for the observed compliance rate differed between treatments. Apparently, for many of the participants in the fifty percent discount treatment, the incentive was not sufficient to induce trial. Moreover, these non-compliers are likely to have attributed to their behavior to a negative disposition toward the newspaper and refused to subscribe when the opportunity was presented. For those accepting the trial offer despite the marginal inducement, compliance was likely to be attributed to their liking of the newspaper, thus enhancing their probability of subscription. On the other hand, the majority of individuals assigned to the premium condition complied with the trial request but were probably uncertain whether their compliance was attributable to a favorable disposition toward the newspaper or to the attractive incentive. In turn, this uncertainty regarding the appropriate cause of their behavior reduced the subscription rate relative to acceptors of the fifty percent discount. If this interpretation is correct, then the failure to observe differences between the premium and fifty percent discount condition is interpretable in self-perception terms.

Although alternative explanations for the similarity in subscription induced by a fifty percent discount and premium offer are available, they appear less tenable than the interpretation offered above. For example, it may be argued that the premium, which was a fifty percent gift certificate at a popular fast food franchise, induced a sufficiently high degree of positive affect toward the newspaper that it obviated any discounting that may have taken place. While this response may have occurred, for some participants it does not explain the aggregate finding. If it did, then one would predict higher rates of compliance among persons accepting the premium offer than among those accepting the fifty percent discount. In fact, just the opposite was found (Table 2).

#### Experiment II: Test of the Self-Perception Mechanism

##### Overview

In Experiment I, qualified support for the self-perception explanation of the foot-in-the-door phenomenon was obtained by manipulating the hypothesized antecedents of self-attribution and observing the effects on behavior. Experiment II sought to clarify further the efficacy of self-perception theory by examining the effects of attributions on attitudes predicted by the theory. More specifically, subjects were asked to read voluntarily a persuasive appeal attributed either to a high or low credibility source and to sign a petition supporting

the communication issue. A series of scaled questions were administered to determine subjects' attitudes toward the issue either before or after they had made a decision regarding whether or not they would sign the petition. In contrast to Experiment I, where incentives served as circumstantial cues, in Experiment II source credibility constituted the factor that mediated subjects' causal attributions. Moreover, Experiment II provided a test of the self-perception explanation of the foot-in-the-door in a situation where a communicator presents information about the attitudinal object in order to gain compliance--a situation that is common in purchase and consumption contexts.

According to self-perception theory, the interaction between source credibility and position of the behavioral request should be significant. Given compliance with the request made by a low credibility source, subjects' attitudes toward the issue should be more positive when determined after they have complied with a request than before compliance has occurred. Observing that they have complied despite the presence of a low credibility source should augment the attribution of behavior to internal causes. If their behavior is not available to subjects as a cue regarding their attitudes, then the situation entails a persuasion paradigm where it has been repeatedly observed that low credibility sources induce relatively little attitude change. On the other hand, if the appeal is attributed to a highly credible source, subjects will be uncertain about the underlying causes of their compliance. It may be attributable to either some personal feeling about the issue (i.e., an internal cause) or to the fact that the message was presented by a highly credible source. As a result, subjects should exhibit a less positive attitude toward the issue in this situation than when the attitudinal posttest precedes behavior.

A similar line of reasoning can be used to predict the attitudes of subjects who did not comply with the behavioral request. Non-compliant subjects exposed to a low credibility communicator are likely to be uncertain whether their failure to comply was due to the low credibility of the source or the fact that they have a negative attitude towards the issue. Conversely, if source credibility is high, non-compliance is more likely to be attributed to subjects' internal dispositions; the requested behavior was not performed despite the fact that source credibility was high.

##### Method

**Subjects.** The 110 participants in the study were recruited from three sections of an introductory course in management. During one of the regular class hours, the students were approached to participate in a study that was being conducted on federal legislation. Participation was made voluntary and the sponsors of the study were described as an external group who had sought the assistance of the course instructor to administer the study. All the students agreed to participate.<sup>3</sup>

**Procedure.** Experimental subjects were given booklets containing the independent and dependent variables by their class instructor. They were asked to read a

<sup>3</sup>An informed consent procedure (Berscheid, Baron, Dernier, and Libman, 1974) administered during the debriefing revealed that less than five percent of the subjects had any objections to the experimental talk.

one-page message which stated the goals of the Consumer Protection Agency Bill (S.707, then pending before the Senate) and which presented arguments for supporting its passage. The credibility of the source was manipulated at the beginning of the communication. For half the subjects randomly assigned to the high credibility condition, the source was described as a Harvard-trained lawyer who was a recognized expert regarding consumer issues. His support for this Bill was made distinctive by describing him as a person who generally did not favor government controls. In contrast, for those subjects assigned to the low credibility condition, the source was portrayed as an individual with no special expertise but who was interested in consumer protection because it represented a job opportunity as a consumer lobbyist. Further, he was described as a person who held socialistic views and favored all government controls. A manipulation check indicated this induction was successful.

The second independent variable involved manipulating the position of the dependent measures. The dependent variables included subjects' attitude toward the Bill, measured by six items on a seven-point Likert-type scale, their evaluation of the source, assessed on the basis of responses to nine semantic differential items, and their behavioral compliance, determined by whether or not they signed a petition to be sent to their Senators urging their representatives to vote for the Bill. For half the subjects, the verbal scales were administered after they had the opportunity to comply with or reject the request to support the Bill. The remaining subjects completed the verbal scales first and then responded to the behavioral request.

#### Results

The two experimental variables did not affect the extent of compliance with the request to sign the petition favoring the Consumer Protection Agency Bill. Of the subjects randomly assigned to the high credibility source condition, 58.18 percent complied, while 43.64 percent of those assigned to the low credibility condition signed the petition ( $Z=.19$ ,  $p>.40$ ). Moreover, whether the request to sign the petition preceded or followed the administration of the attitude measure did not significantly affect compliance. Forty-one percent of the subjects complied when the petition request followed the attitude measures and 61.5 percent complied when it was administered immediately after the communication ( $Z=.19$ ,  $p>.40$ ).

To determine the effects of the experimental variables on subjects' attitudes toward the Consumer Protection Agency Bill, an analysis of variance was performed (Table 4). Consistent with the findings relative to behavioral compliance with the request to sign the petition, source credibility and position of the behavioral request did not have a significant effect on attitudes. The compliance main effect, however, was significant ( $F=52.798$ ,  $d.f.=1/102$ ,  $p<.01$ ); those subjects who had self-selected themselves to sign the petition evinced more favorable attitudes toward the Consumer Protection Agency Bill than those who refused to comply with the petition request.

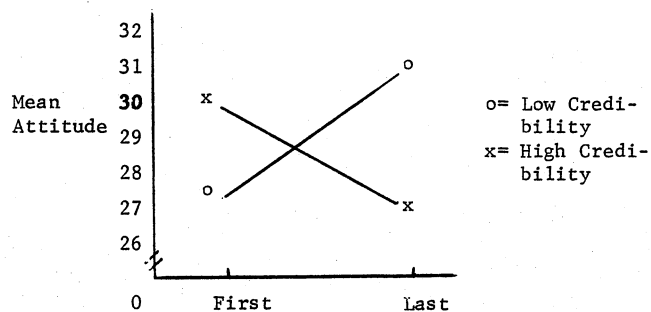
The source X position of behavioral request interaction was also significant ( $F=5.915$ ,  $d.f.=1/102$ ,  $p<.05$ ). This effect is attributable to the fact that the highly credible source induced a more positive attitude toward the Bill when the attitude measure was administered before the behavioral request, whereas the low credibility source was more persuasive if behavioral request preceded the administration of the attitudinal post test (Figure 1). Furthermore, separate analysis of variance for compliers and non-compliers yielded a significant source X position interaction for those who complied

TABLE 4  
Analysis of Variance on Mean Attitudes  
Toward Consumer Protection Bill

Source	d.f.	MS	F
Source Credibility (A)	1	71.657	1.635
Position (B)	1	57.297	1.307**
Behavior (C)	1	2314.378	52.798**
A x B	1	259.285	5.915*
A x C	1	68.810	1.570
B x C	1	.113	<1
A x B x C	1	10.202	<1
Error	102	43.835	

\*\*  $p < .01$   
\*  $p < .05$

FIGURE 1  
Mean Attitude as a Function of Position  
of the Behavioral Request



POSITION OF BEHAVIORAL REQUEST

with the request to sign the petition (Table 5,  $F=4.456$ ,  $d.f.=1/52$ ,  $p<.05$ ). For non-compliers, the joint effect of source credibility and position was similar to that for compliers, but it did not reach conventional levels of statistical significance ( $F=2.628$ ,  $d.f.=1/50$ ,  $p=.10$ ).

TABLE 5  
Analysis of Variance of Attitude  
for Compliers and Non-Compliers

Source	d.f.	MS	F
COMPLIERS			
Source Credibility (A)	1	6.298	<1
Position (B)	1	56.767	2.836*
A x B	1	89.180	4.456*
NON COMPLIERS			
Source Credibility (A)	1	89.984	1.312
Position (B)	1	9.891	<1
A x B	1	180.306	2.628

\*  $p < .05$

#### Discussion

Demonstrations of the foot-in-the-door phenomenon have usually been interpreted in terms of self-perception theory (Bem, 1972). According to this explanation the foot-in-the-door is effective because: "compliance with a small request causes the subject to infer that he has a positive attitude toward cooperating with good causes;

in turn, this positive attitude leads to compliance with the larger request" (Pliner, *et al.*, 1974, p. 18). Experiment II provides further support for the self-perception theory explanation of the foot-in-the-door phenomenon. When behavior served as a cue and an external justification or pressure for that behavior was not present, as is the case in most demonstrations of the foot-in-the-door effect, subjects inferred stronger internal dispositions than when behavior could readily be attributed to an external cause. More specifically, for those people who complied with the high credibility source's request to sign the petition favoring the Consumer Protection Agency Bill, the source's credibility served as a discounting cue. Since compliance could be attributed to both internal dispositions and an external cause (i.e., source credibility), individuals in this treatment exhibited relatively unfavorable attitudes toward the Bill. Conversely, when the communicator was of low credibility, the source cue augmented the attribution of compliant behavior to internal or personal causes. Hence, subjects evinced a relatively favorable attitude toward the Bill.

Support for the self-perception theory would have been more compelling if the source credibility X position interaction for those who did not comply with the behavioral request had reached conventional levels of statistical significance. Nevertheless, treatment differences in attitudinal response were in the predicted direction for individuals exposed to a high credibility communicator. Their attitudes were less favorable when determined after non-compliance than before ( $t=1.32$ ,  $d.f.=21$ ,  $p=.10$ ). A smaller change was observed among individuals receiving the message from a less credible source ( $t=.75$ ,  $d.f.=29$ ,  $p>.40$ ).

The failure to observe a significant source x position interaction among non-compliers could be attributable to the nature of the experimental task. Individuals who complied to the behavioral request may have had fewer alternative explanations for their behavior; either their own attitudes were favorable (when source credibility was low), or the perceived external justification was high (when source credibility was high). Non-compliers, on the other hand, may have had a wider range of reasons for their behavior than did those who signed the petition. For example, they may have had a favorable attitude toward the Bill, but not toward signing the petition; or they may not have wanted to get involved; or they may have had unfavorable attitudes toward the Bill. This may have caused the greater variability in attitudinal responses observed among non-compliers than among compliers ( $t=3.227$ ,  $d.f.53/55$ ,  $p<.01$ ).

Finally, contrary to the self-perception prediction, the differences in attitude induced by the low credibility source before and after the behavior request were not significant either for compliers ( $t=.46$ ,  $d.f.=22$ ,  $p>.40$ ) or for non-compliers ( $t=.75$ ,  $d.f.=29$ ,  $p>.40$ ). This may be attributable to the fact that the low credibility source was not perceived in a very negative light. However, since the present investigation entailed gaining compliance with a behavioral request, the source could be disparaged only to a very limited extent. If a spokesman is seeking compliance, it is unlikely that he would stress biographical data that is particularly damaging to his cause. As a result of this inability to severely disparage the source, the low credibility source may have provided subjects with some external justification for behavior, accounting for the finding that attitudes toward the Bill were not significantly more favorable when behavior preceded rather than followed attitude measurement in the low credibility condition.

From a practical perspective, the experiments reported in this paper suggest that the foot-in-the-door strategy is a viable personal influence strategy providing that the "appropriate" level of incentive is employed to motivate trial. More specifically, some incentive, whether it is operationalized in terms of discounts, source credibility, or some other facilitator, is necessary to induce trial in commercial contexts. However, very substantial incentives may cause individuals to attribute trial to the incentive rather than a positive disposition toward the attitude object. If the incentives subsequently are retracted, the likelihood of compliant behavior diminishes, since the reason for the initial trial is no longer present. Of course, this finding is likely to be maintained only if the product in question is not clearly superior to competitive offerings.

From a theoretical perspective, investigations reported in the paper support the self-perception proposition that behavior, and the circumstances in which it occurs, are important determinants of the attitudes people acquire and that these attitudes, in turn, guide subsequent action. On this basis, it may be hypothesized that people will exhibit attitude-behavior consistency when they have had substantial experience with the attitude object that can be attributed to personal dispositions. Operationally, individuals' confidence in their attitude can serve as a measure of the extent to which behavior has been attributed to a positive disposition toward the object. Indeed, Sample and Warland (1973) observed that factors other than subjects' attitudes did not account for the variance in their behavior when they were highly confident in these judgments.

Finally, the present research provides a basis for constructing a taxonomy of consumption situations. It suggests that situational factors, such as incentives and communicator credibility, can be classified as facilitators of current behavior but inhibitors of subsequent performance, since they undermine the attribution of behavior to internal causes. Other situational variables, such as high price and low credibility, may have the opposite effect. This classification, and self-perception theory on which it is based, question the utility of Belk's (1974) notion of situation which he defines as the "factors particular to a time and place of observation which do not follow from a knowledge of personal ... and stimulus ... attributes, and which have a demonstrable and systematic effect on current behavior" (1974, p. 157). In contrast to Belk's focus on current behavior, the present research indicates the importance of examining the impact of situations on current as well as future behavioral responses.

#### References

- D.A. Aaker, "The Long-Term Value of Temporary Price Reductions," unpublished doctoral dissertation, Stanford University, 1969.
- D.A. Aaker, "A Measure of Brand Acceptance," Journal of Marketing Research, 9(1972), 160-167.
- E. Aronson, J.A. Turner and J.M. Carlsmith, "Communicator Credibility and Communicator Discrepancy as Determinants of Opinion Change," Journal of Abnormal and Social Psychology, 67 (1963), 31-37.

- R.A. Baron, "The 'Foot-in-the-door' Phenomenon: Mediating Effects of Size of First Request and Sex of Requester," Bulletin of Psychonomic Society, 2 (1973), 113-114.
- R. Belk, "An Exploratory Assessment of Situational Effects in Buyer Behavior," Journal of Marketing Research, 11 (1974), 156-163.
- D.J. Bem, "Self-perception Theory," in L. Berkowitz (ed.), Advances in Experimental Social Psychology (New York: Academic Press, 1972).
- E. Berscheid, R.S. Baron, M. Dernier and M. Libman, "Anticipating Informed Consent," American Psychologist, 7 (1973), 913-925.
- B.J. Calder and B.M. Staw, "Self-perception of Intrinsic and Extrinsic Motivation," Journal of Personality and Social Psychology, 31 (1975), 599-605.
- R.B. Cialdini, "A Test of Two Techniques for Inducing Verbal, Behavioral, and Further Compliance with a Request to Give Blood," in review.
- E. L. Deci, "The Effects of Externally Mediated Rewards on Intrinsic Motivation," Journal of Personality and Social Psychology, 18 (1971), 105-115.
- A. N. Doob, J.M. Carlsmith, J. L. Freedman, T.K. Landauer and S. Tom, "Effect of Initial Selling Price on Subsequent Sales," Journal of Personality and Social Psychology, 11 (1969), 345-350.
- J. L. Freedman and S.C. Fraser, "Compliance Without Pressure: The Foot-in-the-door Technique," Journal of Personality and Social Psychology, 4 (1966), 195-202.
- H. H. Kelley, "Attribution in Social Interaction," in E. E. Jones, et al., Attribution: Perceiving the Causes of Behavior (Morristown, N.J.: General Learning Press, 1971), 1-26.
- A.W. Kruglanski, A. Riter, A. Amitai, B. Margolin, L. Shabtai and D. Zaksh, "Can Money Enhance Intrinsic Motivation: A Test of the Content-Consequence Hypothesis," Journal of Personality and Social Psychology, 31 (1975), 744-750.
- M. R. Lepper, and D. Greene, "Turning Play into Work: Effects of Adult Surveillance and Extrinsic Rewards on Children's Intrinsic Motivation," Journal of Personality and Social Psychology, 31 (1975), 479-486.
- L. Marascuilo, "Large Sample Multiple Comparisons," Psychological Bulletin, 65 (1966), 280-290.
- P. Pliner, H. Hart, J. Kohl, and D. Saabi, "Compliance without Pressure: Some Further Data on the Foot-in-the-door Technique," Journal of Experimental Social Psychology, 10 (1974), 17-22.
- J. Sample and R. Warland, "Attitude and the Prediction of Behavior," Social Forces, 51 (1973), 292-304.
- M. Snyder and M.R. Cunningham, "To Comply or Not Comply: Testing the Self-Perception Explanation of the 'Foot-in-the-door' Phenomenon," Journal of Personality and Social Psychology, 31 (1975), 64-67.
- S.W. Uranowitz, "Helping and Self-Attributions: A Field Experiment," Journal of Personality and Social Psychology, 31 (1975), 852-854.



## DYADIC INTERACTION: AN EXCHANGE PROCESS

David T. Wilson, The Pennsylvania State University

Selling may be viewed as a process in which two individuals exchange items of value. In the most simple situation, the buyer receives a product and the seller receives money. In reality, it is likely that buyer and seller exchange attributes with both physical and psychological values.

This exchange concept will be expanded later, but first it is necessary to put it in a context. Exchange does not take place in a vacuum but as part of a buyer-seller interaction process. This paper presents a simple-minded view of this buyer-seller interaction process as background to the central theme of the paper -- dyadic interaction as an exchange process. Throughout the paper the focus has been on developing simple views on critical issues rather than mere elegant models and concepts that may encompass more of reality but provide less understanding.

### The Dyadic Process Model

This model is basically concerned with the development of a long-term buyer-seller relationship rather than "one-shot" selling situations. It begins with an initial meeting between the buyer and seller and moves through a number of stages which presumably take place over time and a number of meetings of the dyad. This does not preclude the adaptation of the model to single contact situations.

Figure 1 depicts the model. A strong similarity exists to the Lavidge Steiner (1961) and AIDA (Strong 1925) models. The duration of the stages overlap each other as a number of activities may be carried on simultaneously. The requirements of the basic stages must be met in order to move to the advanced stages of the model. The different stages are only to suggest that they are the focus of the major efforts of the dyad and the other stages receive less effort during the current stage. The model proposed here attempts to reflect the influence of some of the sales research conducted in recent years.

There are two basic assumptions underlying the model:

1. The buyer is attempting to secure a bundle of attributes, both tangible and psychological, from the seller. These attributes may be related to the product, to the company and to the salesman. To illustrate, the product may provide performance attributes that are highly valued by the buyer, the company provides general reputation attributes and the salesman may provide personal service and reliability attributes.
2. The buyer-seller relationship develops over a period of time. Examples of this type of relationship are numerous in organizational buying situations, personal estate planning, life insurance situations and in some retail situations where strong customer loyalties are developed.

In addition, it is recognized that in many organizational and family buying situations, a buying center exists that involves more than one individual. Once recognizing this buying center, it will be ignored for the sake of simplicity in developing the process model. The model can

also be expanded to include goals of each individual at each stage and influences upon the individuals at each stage. These complications will also be ignored as complexity at this embryonic stage can be lethal.

### Source Legitimization

The model assumes the buyer and seller come together in something akin to the Howard and Sheth (1969) extensive problem-solving situation. Monoky, Mathews and Wilson (1975) found that sources of information are differentially preferred as a function of the type of selling situation. The situations ranged from a new task to a rebuy situation. This data would suggest that the salesman's role shifts over time as the buying situation moves from the new task to the rebuy situation.

In the new task situation the salesperson needs to develop source credibility and legitimization. Unless this basic acceptability is developed, further communication is likely to become quite ineffective if not impossible.

Although the salesperson may be perceived on a number of dimensions, the dimensions of expert-non expert and similar-dissimilar have received the greatest attention in the literature.

Based on current studies, it seems likely that in this early stage of the sales process that the salesperson attribute of "expert" may dominate the attribute of "similarity." Woodside and Davenport (1974) in their study of an extensive problem-solving retail sales situation offer some support to this notion. Busch and Wilson (1975) in a simulated sales situation found expert attributes in salesmen to be quite powerful in achieving attitude and behavioral changes in buyers. Their study also concluded that the expert was more powerful than the referent in the early stage. It should be noted that both of these experimental situations were "one-shot" sales situations and as such there may be a bias toward the expert attribute. It seems reasonable that as the buyer-seller relationship matures, the attribute of similarity may take on more value.

The basic objective of this first stage is to establish the salesman as a legitimate and credible partner in the dyadic interaction process. It is likely that this process continues through the early interactions of the dyad.

### Information Exchange - Problem Identification

This stage involves bounding the problem to be solved through a purchase. The amount and nature of the information exchange is likely a function of the relationship established in the previous stage. A salesperson perceived as an expert and similar is likely to be given more information quicker than other salespersons. All salespersons may get the standard type of information but the more favored salesperson may get additional data such as which attributes are important and who is the key decision maker in the buying center.

The salesperson attempts to establish the nature of the problem in order to be able to suggest an attribute package that will result in a sale. The salesperson is also concerned about positioning the attribute package with

respect to competition and in developing a strong bargaining base.

#### Attribute Delineation

In this stage, the dyad develops the bundle of attributes that will be exchanged. Many of the attributes will be explicitly discussed; for example, product features and credit terms, while other attributes may be implicitly determined. Olshavsky (1973) observed that the salesman clearly influenced attribute determination and evaluation. Again, this study is limited by its one-shot nature in a retail setting where the buyer may lack the buying expertise that is present in organizational purchasing. Nevertheless, the study does indicate that although both parties developed the attribute set, the salesman can have a major influence in the development of the attribute set through guidance and direction of the buyer.

It seems reasonable to expect the determinant attributes to shift over time. Attributes such as product quality, delivery, etc., may remain important but not determinant in how the dyad relationship is maintained. Interpersonal exchanges between the buyer and seller may become more important as the purchasing problem moves to the rebuy situation.

Also of importance is the nature of the product. If there are clear product differences they may be used to differentiate the various attribute packages. However, if substantial product differences do not exist, then intangible attributes may dominate.

#### Attribute Value Negotiation

The determining of the attribute set and the exchange rate of each attribute can be viewed as a bargaining process. Pennington (1968) p. 262, concluded "that bargaining behavior between customers and salespersons, although not overwhelming in frequency or volumes, exerts an important impact on ultimate purchase outcomes." He found that the greatest bargaining was "on the presentation of concession limits -- such as presentation of prices, delivery dates, product features -- by the salesman, and statements of desired price ranges, styles, and product features by the customers," (p. 257). These findings support the notion of bargaining the rate of exchange on attributes. The value of individual attributes may be subject to limited negotiation. The seller will likely attempt to increase the importance in the choice of his strong attributes while the buyer may attempt to suggest an increased importance of the seller's weaker attributes in hopes of gaining an increased level of other attributes. For example, the buyer may suggest that price is important when he really hopes to gain better credit terms. In other words, there may be a trading of the amount of each attribute that will make up the final package of attributes.

Wilson, Mathews and Monoky (1973) and Mathews, Wilson and Monoky (1972) found that perceived similarity improved cooperation in a simulated buyer-seller bargaining situation. A "prisoner's dilemma" game was used as the bargaining vehicle. They (1973) used an extended Fishbein attitude model and found that although attitude towards the act of cooperation was important, the normative component was influential for the similar subjects and not influential for the dissimilar subjects.

The similarity variable may be operational throughout the total process. Its power to influence behavior may be greatest at the bargaining stage.

#### Relationship Maintenance

In the final stage of the model, the dyad maintains and builds upon their relationships. New personal attributes of the relationship may develop which enables it to grow from a business relationship to a more personal business/friendship relationship. A certain amount of implicit bargaining over exchange values may take place particularly if problems with performance attributes arises. From a salesperson's point of view, the maintenance of friendly accounts is much easier than the development of new accounts.

#### Dyadic Interaction: An Exchange Process

The development of the process model seems a roundabout road to the focus of this paper which is examining dyadic seller behavior as an exchange process. It was necessary to establish the notion that this dyadic relationship goes through a number of stages during which a bundle of attributes valuable to the buyer is developed.

The basic premise of the concept of an exchange process derives from Homan's (1961, p. 7) contention that "elementary social behavior is the face-to-face contact between individuals, in which the reward each gets from the behavior of the other is relatively direct and immediate." He suggests that this exchange of interaction can yield a profit to both parties if you "give the other man behavior that is more valuable to him than it is costly to you and to get from him behavior that is more valuable to you than it is costly to him" (p. 62).

Homan's concept may be extrapolated to the buyer-seller dyad as the notion that each individual develops a bundle of attributes that can be exchanged and which have utility to him. The successful dyad exchanges these bundles of attributes at a profit. On one level, the salesman provides product and receives payment while on another level he may provide advice and receive esteem. The buyer receives product and gives payment. He also may receive this advice which helps him solve a purchasing problem and gives esteem as payment. The exchange of advice may not directly relate to the salesman's product.

Again to simplify the discussion, the linkage of the value of the attribute to the buyer or seller may depend upon the attribute to satisfy the needs of others within either the buyer's or seller's firms will be recognized and then ignored. The exploration of these linkages needs a paper in itself.

To illustrate the exchange concept, assume a purchasing situation in which the buyer must choose between the offerings of three salespersons. Each salesperson goes through the early stages of the process model of Figure 1; i.e., source legitimization, information exchange, and attribute delimitation. The individual salesperson attempts to establish a set of attributes that have value to the buyer and if possible, are unique from his competitor's offerings. Hence, the buyer is confronted with three unique sets of attributes ( $X_{nj}$ ). The subscript  $j$  refers to the salesman, the subscript  $n$  refers to the attributes, and the subscript  $q$  refers to the level of attributes in the set. The buyer must select  $X_{nq1}$ ,  $X_{nq2}$ , or  $X_{nq3}$ . It is assumed that each of these attribute sets have some cost to the buyer and that the buyer seeks to maximize his profit in the transaction, that is, the buyer seeks the maximum multi-attribute utility (MAU). The choice is now to select from the utility set that  $U_{nj}$  with the highest value. It is assumed that individual attributes can be positive or negative and that the summing of the attribute utility nets out to the "profit" of the relationship.

Profit in this situation is a nebulous concept but may be thought of as the netting out of positive attributes and negative attributes. For example, could price be a negative attribute. It is suspected that exploratory research may eliminate the need for the profit notion and that a summation of the values positive and negative attributes will be sufficient.

Recalling that the attribute sets contain product, company and salesman attributes it is then possible to examine the impact of the salesman upon the total value of the attribute package.

The exchange notion enters in that the salesman must allocate his time and effort over a number of buyers and he presumably seeks to maximize his total return by choosing to deal with those buyers who offer him a "profit." Therefore, the buyer must present a set of attributes to the salesman, the principle one being the monetary value of the order. This concept may become even more relevant in this era of shortages and sellers' markets.

#### Measurement of the Value Attribute Sets

There has been a considerable amount of research conducted on MAU models, both in the laboratory and in the field. Huber (1974a, 1974b) presents a useful summary of the state of the art. The attribute set in the selling dyad is likely to be much more complex than those that have been reported in the literature.

A general additive form of the multi-attribute utility model is

$$U(x) = \sum_{n=1}^n b_n U(X_{nqj})$$

where  $b_n$  is a measure of the relative importance of the attributes and can be represented on a 0 to 1.0 scale.  $U(X_{nq})$  is the utility value of attribute  $X_n$  at the  $q$  level. It is suggested that since the attribute sets include salesman attributes which are unique to the individual, it might be useful at this early stage of development to describe the utility sets as  $U(X_{nqj})$

where  $j$  refers to the salesman. If in fact the attribute sets prove to be consistent across salesmen, then only the level of an attribute need be used. If uncertainty associated with the attributes is considered a subjective expected multi-attribute model (SEMAU) is created. Wilson and Busch (1973) in a test of the association between an attitude intention model and a SEU model found a high correlation between the predictions of the two models. Although the models are theoretically different, their robustness may allow them to approximate the same results.

Huber (1975) had similar results when he attempted to use different models to predict individual preferences on stimuli defined by physical characteristics. He suggests that the robustness of his models allows them to approximate preference surfaces in spite of theoretical invalidity," (p. 295).

The large literature on expectation models may provide some guidance in developing empirical tests of the MAU proposition.

#### Conclusions

This paper has attempted to describe one way of conceptualizing buyer-seller dyadic interaction. The focus

has been on the changing relationship over time as the dyad goes through the stages of the selling process model. In particular the development of a bundle of attributes that has utility to the buyer and a similar bundle that has utility to the seller was examined as a basis for an exchange process.

The importance of individual attributes likely varies as a function of the stage of the relationship. Careful work needs to be done on the stages in which the attributes set is developed and values of the individual attributes are negotiated.

On the seller side insights might be gained by studying the MAU sets that a salesman perceives in his customer and prospect set. How does he assign his time? To what extent does he interact with buyers where he receives a "profit" or has the lowest cost? The exchange process, particularly involving non-product attributes in situations where product and price is relatively homogeneous, for example, the steel industry is worth study. Then, if the researcher is astute enough, it might be possible to describe and measure the psychological attributes that are operating.

Obviously, this whole paper is a beginning attempt to conceptualize a difficult process. It is likely that many changes will be made as empirical studies provide additional information. The task now is to begin a series of studies that will help to clarify the area.

Figure 1

#### A DYADIC SALES PROCESS MODEL

Source Legitimization

Info. Exchange  
Problem Iden.

Attribute  
Delineation

Attribute  
Value

Negotiation

Relationship  
Maintenance

#### REFERENCES

- Paul Busch and David T. Wilson, "An Experimental Analysis of a Salesman's Expert and Referent Bases of Social Power in the Buyer-Seller Dyad," unpublished manuscript.
- George C. Homans, Social Behavior: Its Elementary Forms. New York: Harcourt, Brace and World, Inc., 1961.
- John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior. New York: John Wiley, 1969.
- George P. Huber, "Methods for Qualifying Subjective Probabilities and Multi-Attribute Utilities," Decision Sciences, Vol. 5, 1974.
- George P. Huber, "Multi-Attribute Utility Models: A Review of Field and Field-Like Studies," Management Science, Vol. 20, No. 10 (June, 1974).
- Joel Huber, "Predicting Preferences on Experimental Bundles of Attributes: A Comparison of Models," Journal of Marketing Research, Vol. XII, 290-7 (August, 1975).

Robert J. Lavidge and Gary A. Steiner, "A Model for Predictive Measurements of Advertising Effectiveness," Journal of Marketing, 61 (October, 1961).

H. Lee Mathews, David T. Wilson, and John F. Monoky, Jr., "Bargaining Behavior in a Buyer-Seller Dyad," Journal of Marketing Research, Vol. IX, 103-5, (February, 1972).

John F. Monoky, Jr., H. Lee Mathews, and David T. Wilson, "Information Source Preference by Industrial Buyers as a Function of the Buying Situation," Pennsylvania State University Working Series.

Richard W. Olshavsky, "Customer-Salesman Interaction in Appliance Retailing," Journal of Marketing Research, Vol. X, 208-12, (May, 1973).

Allan L. Pennington, "Customer-Salesman Bargaining Behavior in Retail Transactions," Journal of Marketing Research, Vol. V, 255-62, (August, 1968).

E. K. Strong, The Psychology of Selling. New York: McGraw-Hill Co., 1925.

David T. Wilson and Paul Busch, "A Theoretically Integrative Approach to the Study of Decision Making: An Empirical Analysis," Proceedings of the Fifth Annual Conference, American Institute for Decision Sciences, November, 1973.

David T. Wilson, H. Lee Mathews, and John F. Monoky, Jr., "Attitude as a Predictor of Behavior in a Buyer-Seller Bargaining Situation: An Experimental Approach," The Proceedings of the American Marketing Association, August, 1973.

Arch G. Woodside and J. William Davenport, Jr., "The Effect of Salesman Similarity on Consumer Purchasing Behavior," Journal of Marketing Research, Vol. XI, 198-202, (May, 1974).

## CONSUMER RESPONSE TO ALTERNATIVE SELLING STRATEGIES: A FIELD EXPERIMENT

Arch G. Woodside, University of South Carolina  
Robert E. Pitts (student), University of South Carolina

### Abstract

The analyses of the separate and joint effects on sales of salesman perceived expertise, price, and the absence or presence of customer purchase pals are reported. The three main effects were statistically significant: sales increased with increases in salesman perceived expertise, the presence of purchase pals, and price increases (\$1.98 to \$3.98). The expertise-by-price interaction was statistically significant using multiple regression and multivariate probit analyses and four price treatments. The purchase pal-by-price interaction was statistically significant using three price treatments. The need for hypothesizing and empirically testing interaction effects on demand of marketing variables is emphasized.

### Introduction

Marketing decisions must be made in the context of insufficient information about processes that are dynamic, nonlinear, lagged, stochastic, interactive, and downright difficult (Kotler, 1972, p. xi).

Despite Kotler's famous quote of 1967 and 1972, few empirical studies exist on interaction and other effects on sales of pricing, product, promotion, and distribution levels which are downright difficult to explain or measure. Recently, Curhan (1974) and McCann (1974) used fractional factorial analysis of variance and regression analysis respectively to measure the main and interaction effects of multiple marketing decision variables; unfortunately, the discussions of their findings focused nearly exclusively on the main effects of the decision variables.

Kotler (1971) maintains that the market's response to variations in the level of any one marketing input is conditional on the level of the other decision variables. Furthermore, the variation of two or more marketing activities at the same time can have synergistic effects that are greater or less than the sum of the separate effects.

Some progress has been made on developing models of interaction effects on a conceptual level and measuring them on an empirical level (e.g., Kotler, 1964; Green, 1973).

This article reports on the analyses of the separate and joint effects on sales of two marketing decision variables and one consumer related variable. The analyses are based on a field experiment using two levels of a promotional variable, salesman expertise, and four price levels of a consumer semi-durable product. The presence or absence of other persons, i.e., purchase pals, during the sales presentations was the consumer-related variable in the study.

### Hypotheses

The general form of the hypothesized model of the effects on sales of salesman expertise, price, and purchase pal includes main and interaction effect terms:

$$Q = a + b_1E + b_2Pr + b_3P + b_4(EPr) + b_5(EP) + b_6(PrP) + b_7(EPrP) \quad (1)$$

where: Q = product sales  
E = expertise  
Pr = price  
P = pal  
 $b_i$  = regression coefficients ( $\beta$  = standardized partial regression coefficients).

Intuitively, a positive main effect of perceived expertise attached to the salesman and sales while a negative main effect of price and sales would be expected.

- H<sub>1</sub>: An increase in the level of perceived expertise attached to the salesman produces a greater likelihood of purchase by the customer.  
H<sub>2</sub>: An increase in the level of price produces a decrease in likelihood of purchase by the customer.

As a rationale for H<sub>1</sub>, Woodside and Davenport (1974) have used Kelman's (1961, 1965) conceptualization of source credibility and internalization of the message by the receiver: the greater the communicator's perceived credibility by the recipient, the greater the likelihood that the recipient will accept the influencing message because it is congruent with the recipient's value system, i.e., internalization has occurred. Brock (1965) has also found that credibility of the source affected decisions made by customers (change from a customer's selected brand of paint to one advocated by the salesman).

The law of demand as expressed in H<sub>2</sub> is open to challenge depending upon other factors operating in the market. Four possible reasons for expecting a positive price-demand relationship have been reviewed by Kotler (1971):

1. With the greater purchasing power following a substantial price reduction of a product purchased in substantial quantities, consumers may decide to switch to more expensive products, i.e., the income effect.
2. Higher price is likely to increase demand for certain goods because of the phenomenon of "conspicuous consumption," i.e., the Veblen effect.
3. An expectation effect: when a price reduction is seen as the beginning of a possible wave of further price reductions, many buyers may withhold their purchases.
4. The quality effect: in situations in which buyers are not well informed about the respective merits of competing products, and some risk is involved, buyers may take the price as an indication of quality.

Positive price and perceived-quality relationships have been reported in the marketing literature in situations involving low consumer knowledge and some risk. In a retail experiment, Woodside and Sims (1974) also report a positive price and quantity demanded relationship for a new product, an electric lunch box.

Intuitively, the presence of other persons accompanying a customer would produce a decrease in the effectiveness of a sales presentation, i.e., a decrease in the likelihood of purchase by the customer. A purchase pal may be viewed as the actor with the least at stake in the transaction, offering the customer a more unbiased view of the situation than the salesman; and therefore, the customer may be less affected by the sales presentation when a purchase pal is present since the pal may provide reasons for not purchasing the product.

H<sub>3</sub>: The presence of a purchase pal decreases the likelihood of purchase by customers following a sales presentation.

Some support of this hypothesis is provided by Bell's (1967) report that purchase pals (friends of relatives accompanying the customer) made automobile selling more difficult and less pleasant for the salesman.

H<sub>4</sub>: Decreases will be greater in the likelihood of purchase produced by increases in prices, the lower the level of perceived expertise of the salesman.

H<sub>5</sub>: Increases will be greater in the likelihood of purchase produced by increases in expertise if no purchase pal is accompanying the customer compared with increases when a purchase pal is present.

H<sub>6</sub>: Decreases will be greater in the likelihood of purchase produced by increases in prices if a purchase pal is accompanying the customer compared with decreases when no purchase pal is present.

H<sub>7</sub>: The presence of a purchase pal increases the changes in likelihood of purchase produced by price under low versus high perceived salesman expertise conditions.

Hypotheses 4 through 7 are interaction predictions of the three independent variable effects on sales. For specific price increases, the salesman of high versus low perceived expertise could decrease the customer's perceived risk in evaluating the product and consequently price changes should have a smaller impact on sales. This is the rationale for H<sub>4</sub>.

Increasing reliance in the salesman under high versus low expertise conditions is likely to occur when no purchase pal is present. If the customer is not able to seek help from others for making the purchase decision, the resulting internalization effect achieved by increasing the perceived expertise of the salesman has a greater likelihood of occurring. This is the rationale for H<sub>5</sub>.

Higher product prices may produce more comments by purchase pals of the economic risk in purchasing the product, thus the presence of purchase pals should increase the effect of price changes on sales, i.e., produce a greater price elasticity (H<sub>6</sub>).

The presence of a purchase pal provides the customer with an additional source of information which could increase the customer's perception of the product's price; consequently, changes in the likelihood of purchase produced by different prices for high versus low levels of salesman expertise are likely to be greater when purchase pals are present.

The seven hypotheses are shown graphically in Figures 1 through 4. The interaction hypotheses can be stated for the demand curve in the figures:

H<sub>4</sub>:  $[(D_a - D_b) / (Pr_1 - Pr_2)] < [(D_c - D_d) / (Pr_1 - Pr_2)]$   
from Figure 1.

FIGURE 1

Hypothesized Effects of Price and Salesman Expertise

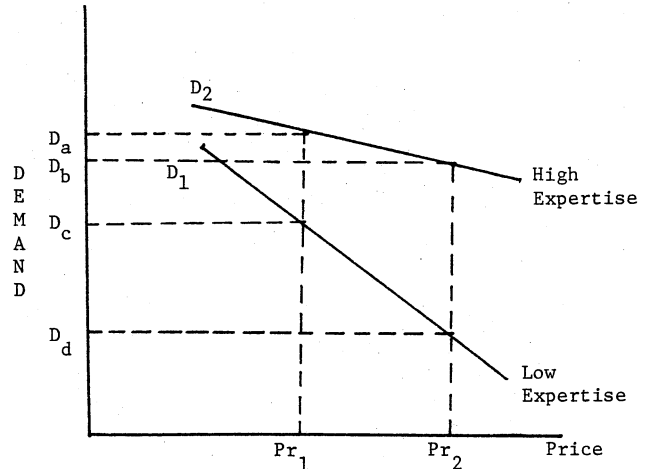


FIGURE 2

Hypothesized Effects of Expertise and Purchase Pals

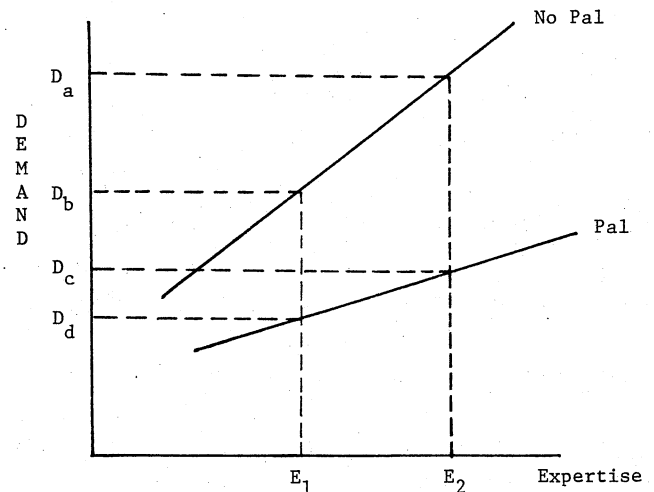


FIGURE 3

Hypothesized Effects of Price and Purchase Pal

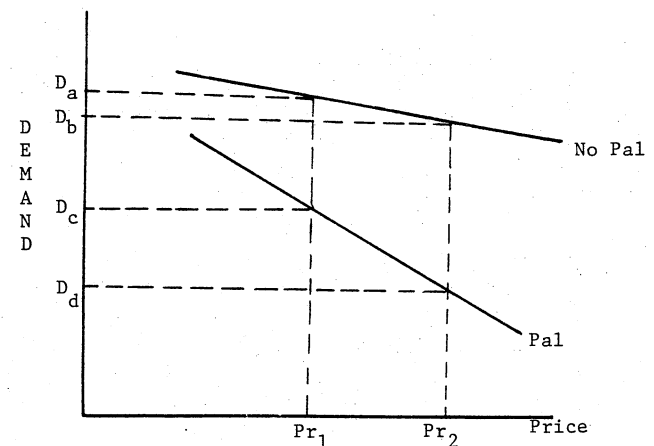
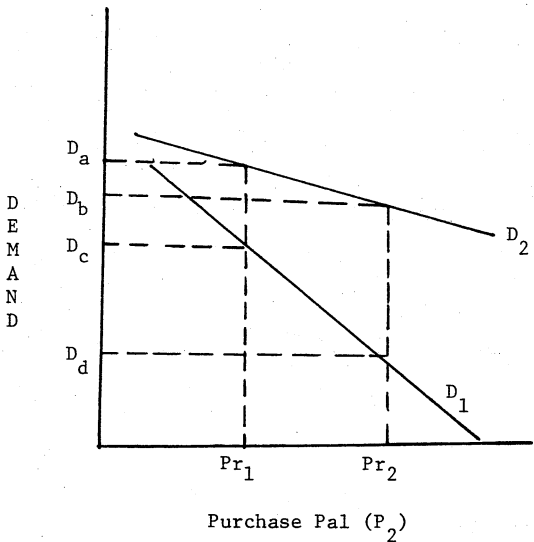
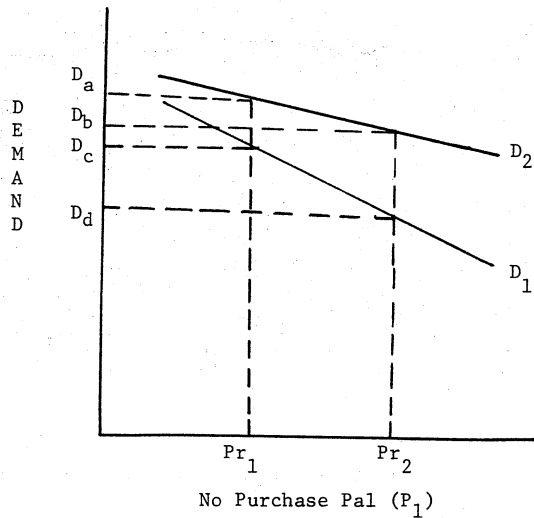


FIGURE 4

Hypothesized Effects of Price, Expertise, and Purchase Pal



- H<sub>5</sub>:  $[(D_a - D_b)/(E_2 - E_1)] > [(D_c - D_d)/(E_2 - E_1)]$   
 from Figure 2.  
 H<sub>6</sub>:  $[(D_a - D_b)/(Pr_1 - Pr_2)] < [(D_a - D_b)/(Pr_1 - Pr_2)]$   
 from Figure 3.  
 H<sub>7</sub>:  $\{[(D_a - D_b)/(Pr_1 - Pr_2)] - [(D_c - D_d)/(Pr_1 - Pr_2)]\}/P_1 <$   
 $\{[(D_a - D_b)/(Pr_1 - Pr_2)] - [(D_c - D_d)/(Pr_1 - Pr_2)]\}/P_2$   
 from Figure 4.

Method

The data of this study are taken from a field experiment set in a retail store from which some results have previously been reported (Woodside and Davenport, 1975). The research design used included two levels of salesman-perceived expertise and four prices which enabled sales comparisons to be made between treatment levels and across treatments. In addition, the salesperson recorded the presence or absence of a purchase

pal after the customer left the store. A control group of subjects received no salesman presentation with the product priced at \$1.98.

A salesperson attempted to induce purchase of a new product innovation among consumers shopping for 8-track stereo music tapes in a small music store in Augusta, Georgia. The salesperson was a woman in her late thirties who was the communicator for all eight treatment conditions. The salesperson was a high school graduate and had not been trained in marketing theory or sales psychology. She was not informed of the predicted results of the study and no additional compensation was offered to her.

Product

The product was the "HC-2001, Head and Capstan Cleaner Kit," manufactured by Becht Electronics, Burbank, California. The kit included two felt pads, head cleaning solutions, and cartridge to be used to clean 8-track players.

The product had been introduced to the market during the month of the experiment and none of the firm's six competitors carried the product. The six competitors were visited during the course of the study by a shopper wanting to buy "some type of cleaning stuff or kit" for his tape player. Two competitors had knowledge of such devices but stated that demand was not great enough to carry them.

The product was somewhat technically complex and its safe use was assumed to be important to a user since the product had to be connected to the tape player. The kit was mounted on cardboard, enclosed in a plastic container, and had a suggested retail price of \$1.98. Operating instructions were printed on one side of the cardboard.

Procedure

The salesperson attempted to induce customers, who had just purchased one or more tapes, to make an additional purchase of the cleaning kit. Selected customers were randomly assigned to one of the eight treatment conditions and to the control group. Treatment conditions were typed and copies placed below the cash register in the store after the copies were randomly mixed. Space was available on the copies to record purchase information. Blank copies represented the control group assignment.

Customers examining 8-track tapes were selected unobtrusively as subjects in the experiment. While a selected subject was examining the tapes, the salesperson looked at the top treatment copy under the register and administered that particular treatment. A display box containing 20 of the cleaning kits was placed near the cash register throughout the experiment. A 6" by 6" card appeared on the display box. The words "8-Track Tape Cleaner Kit" and the price were placed on the card. Different cards were used for the different price treatments. Prices were not listed on the cleaning kits. Purchase response was recorded immediately after the customer left the store. The copy of the treatment administered remained in its top position until the next subject was selected. The salesperson removed the previous customer's treatment copy at this time, noticed her role for the new subject, changed the display card if necessary, and administered the treatment when the subject approached the cash register.

The treatments consisted of the eight combinations of the four price levels and the two expert - non-expert levels. A total of 30 customers were assigned to each combination and to the control group. Purchase behavior was the only dependent variable measured in the study.

The following price levels were included in the study: \$1.98, \$2.98, \$3.98, and \$5.98. This wide range in prices was used to attempt to insure for some significant price effects in the experiment. Hunch lead to the belief that demand would not be affected for a wide range of price. The product's price for the control group was \$1.98.

The salesperson expressed prior purchase by herself of the musical tapes bought by the customer in all treatment conditions, except for the control group. Levels of perceived expertise were defined as the salesperson's oral instructions on how to operate the tape cleaner versus her expressed inability to operate the cleaner.

The salesperson's presentations to the customers were independently observed by one of the experimentors. The prior decision was made to discard the data for the first ten customers given the presentations. A number of small operational problems were corrected in these initial presentations and the salesperson became proficient in delivering the sales messages.

Customers in the control group did not receive any sales presentation but could purchase the product from the display box on the counter near the cash register. Any questions asked about the product by these customers were answered by the salesperson. No customer was rejected for requesting information. The salesperson responded to questions in the treatment role as required. The salesperson recorded whether or not the customer requested additional information. Few customers requested further information.

Store policy was to accept cash, check, or charge card and no discrimination was made for method of payment.

Further details of the field experiment procedure and specific wordings of the expertise conditions are given by Woodside and Davenport (1975). Different operational definitions of expertise could have been used in the experiment which implies theoretical and managerial constraints on examining the findings.

#### Pre-Test of Appeals

Perceived differences of the communicator's message for the expert and non-expert treatments were examined in a pre-test. Students in one class at the University of South Carolina rated the salesperson using the messages on the following qualities for 7-point Likert scales: cold-warm, non-expert - expert, makes me not want to buy-makes me want to buy, and salesperson not familiar with my needs-salesperson is familiar with my needs.

The written appeal given the students began with the following:

Assume you have just purchased some 8-track stereo tapes, and as you are paying for them, the salesperson says:

The students were randomly assigned to receive one of the treatments by randomly mixing the copies of the appeals. The students were told that answering the questions was not connected with the course, not to place their names on the paper, and their questions would be answered afterwards. The \$1.98 price was

used for both treatments in the pre-test.

All results were either statistically significant ( $p < .05$ , t-tests) or supported preconceived hypotheses in directions of mean differences from scores on the Likert scales. The salesperson in the expert treatment was statistically significantly rated more expert, makes me want to buy, and familiar with my needs compared with the salesperson in the non-expert treatment. Further details of the pre-test are discussed in Woodside and Davenport (1974).

#### Results and Discussion

The complete set of data of the experiment are presented in Table 1. The three main effects of expertise, price, and purchase pal were statistically significant. Results are shown in Table 2 of chi-square analyses of these main effects.

TABLE 1

Customer Purchase Behavior for Four Price, Two Salesman Expertise, and Two Purchase Pal Conditions

Price	Salesman Expertise	Purchase Pal	Purchase	No Purchase	n
\$1.98	Expert	Yes	13	2	15
		No	11	4	15
	Nonexpert	Yes	5	7	12
		No	4	14	18
\$2.98	Expert	Yes	9	0	9
		No	12	9	21
	Nonexpert	Yes	6	6	12
		No	5	13	18
\$3.98	Expert	Yes	12	1	13
		No	8	9	17
	Nonexpert	Yes	3	3	6
		No	9	15	24
\$4.98	Expert	Yes	2	4	6
		No	2	22	24
	Nonexpert	Yes	1	9	10
		No	0	20	20
\$1.98	Control		4	26	30

TABLE 2

Main Effects of Price, Salesman, Expertise and Purchase Pal on Customer Purchase Behavior, in Percent

Variable	Level	Purchase	No Purchase	n
Price	\$1.98	55.0%	45.0%	60
	2.98	53.3	46.7	60
	3.98	53.3	46.7	60
	5.98	8.3	91.7	60
Expertise	Expert	57.5	42.5	120
	Nonexpert	27.5	72.5	120
Purchase Pal	Present	61.4	38.6	83
	Absent	32.5	67.5	157
Control	No Treatment and \$1.98 Price	13.3	86.7	30

The first two hypotheses were supported by the analysis. The increase in perceived salesman expertise produced a substantial increase in sales (30%). Only 8.3% of the



customers in the \$5.98 price treatment purchased the product, while 55% of the customers in the \$1.98 treatment purchased the product. However, the differences in percent of product purchases between the \$1.98, \$2.98, and \$3.98 treatments were slight. More than 50% of customers in these treatments purchased the product.

Counter to the third hypothesis, a significant positive relationship between purchase pals and customer purchase behavior was found. Over 61% of the customers shopping with purchase pals purchased the product versus 32.5% purchases when the purchase pal was not present. Two post hoc explanations of this finding can be suggested. A purchase pal may provide an immediate means for social validation (Howard, 1965) that the purchase is a correct choice and thereby reduces the perceived risk in the act of purchase and increases the likelihood of purchase. Secondly, the presence of a purchase pal provides the customer with a more conspicuous role to play in deciding whether or not to purchase the product; thus, the customer may experience a greater need to provide reasons for not purchasing the product. Consequently, the purchase of the product may be the easier action to take for the customer.

Multiple regression analysis and multivariate profit analysis were used to test the significance of the interaction terms as well as the terms for the main effects in equation 1. A quadratic term for price ( $Pr^2$ ) was included in the equation because a curvilinear relationship was believed to likely result from the wide range of prices used in the experiment (\$5.98 - \$1.98).

The dichotomous dependent variable was nonpurchase versus purchase. Effect coding (-1, +1) was used on levels of expertise and no purchase pal versus purchase pal. Effect coding is similar to dummy coding except that the former generates regression coefficients which reflects the linear model. The standardized partial regression coefficients of the independent variables produced using effect coding are equal to the correlation coefficients (phi-coefficients) of the independent and dependent variables, if the independent variables are orthogonal (Woodside, Pitts, and Gewirtz, 1975).

Multivariate probit analysis was used to overcome some of the limitations of using multiple regression analysis on a dichotomous or ordinal level dependent variable. Probit analysis assumes a linear effect on each independent variable as well as a series of break points between categories for the dependent variable. Maximum likelihood estimators are found for these coefficients, along with their asymptotic sampling distributions, as well as an analogue of  $R^2$  is defined to measure goodness of fit (McKelvey and Zavoina, 1975; Kau and Hill, 1972).

McKelvey and Zavoina (1975), and Bettman (1974) review some of the problems of using multiple regression analysis with a dichotomous dependent variable: heteroscedastic disturbances, expected value of the error term greater than zero, and non-normally distributed error term. However, multiple regression is quite robust and departures from its underlying assumptions are unlikely to produce substantially different conclusions in the comparison of the relative importance of the independent variables. The partial regression coefficients and  $R^2$  estimated by multiple regression are more conservative than the coefficients of the coefficients and  $R^2$  produced by probit analysis (McKelvey and Zavoina, 1975).

The resulting equations from multiple regression and probit analyses are shown in Tables 3 and 4. The

TABLE 3  
Analysis of "Best" Model with  
Four Prices<sup>1</sup>

Variables	Regression Analysis			$R^2$ (F) <sup>a</sup>
	b	$\beta$	t	
Price	.308	.921	2.50	.32 (22.07)
Expertise	.307	.621	4.24	
Pal	.119	.229	4.19	
Price <sup>2</sup>	-.051	-1.261	-3.42	
Expertise by Price	-.043	-.349	-2.38	
Constant	.858			

Variables	Probit Analysis			$\hat{R}^2$ ( $\lambda^*$ ) <sup>b</sup>
	b	$\beta$	Z	
Price	1.31	3.385	2.64	.47 (89)
Expertise	.763	1.544	2.84	
Pal	.401	.773	4.03	
Price <sup>2</sup>	-.188	-4.604	-3.46	
Expertise by Price	-.083	-.676	-1.13	
Constant	-1.348			

<sup>a</sup>d.f. = 5,234.

<sup>b</sup> $\lambda^*$  is  $X^2$  with d.f. = 5.

<sup>1</sup>"Best" in that the partial coefficients were each statistically significant and the independent variables each substantially contributed to explaining variance in the dependent variable.

TABLE 4  
Analysis of "Best" Model with  
Three Prices<sup>1</sup>

Variables	Regression Analysis			$R^2$ (F) <sup>a</sup>
	b	$\beta$	t	
Expertise	.290	.582	3.96	.21 (15.67)
Price by Pal	.044	.261	3.86	
Expertise by Price <sup>2</sup>	-.012	-.268	-1.82	
Constant	.576			

Variables	Probit Analysis			$\hat{R}^2$ ( $\lambda^*$ ) <sup>b</sup>
	b	$\beta$	Z	
Expertise	.825	1.657	3.64	.30 (41)
Price by Pal	.130	.773	3.72	
Expertise by Price <sup>2</sup>	-.036	-.769	-1.72	
Constant	.245			

<sup>a</sup>d.f. = 3,176.

<sup>b</sup> $\lambda^*$  is  $X^2$  with d.f. = 3.

<sup>1</sup>See footnote to Table 3.

equations in Table 3 are based on data of all four prices while data for \$1.98 to \$3.98 were used to compute the equations in Table 4. A stepwise solution was used to enter the variables in the multiple regression equations in Tables 3 and 4.

The regression equation in Table 3 produced an  $R^2$  of .32 ( $F = 22.07$ ,  $d.f. = 5, 234$ ,  $p < .001$ ). The  $\hat{R}^2 = .47$  following probit analysis was also significant with the test statistic distributed as  $X^2$  ( $\lambda^* = 89$ ,  $d.f. = 5$ ,  $p < .001$ ).

Coefficients of the three main effects of price, expertise, and purchase pal were each statistically significant ( $p < .01$ , two-tailed tests) in both the regression and probit equations in Table 3. The coefficients of price-squared was also statistically significant ( $p < .02$ ). The expertise-by-price interaction coefficient was the only interaction coefficient statistically significant ( $p < .02$  from the regression coefficient).

No other variables were statistically significantly related to purchase ( $p < .10$ ) based upon analysis of variance of the data.

$R^2$  and  $\hat{R}^2$  in Table 4 are also statistically significant ( $p < .001$ ). The standardized partial regression coefficients with price instead of price-squared are similar to the coefficients found in Table 4:

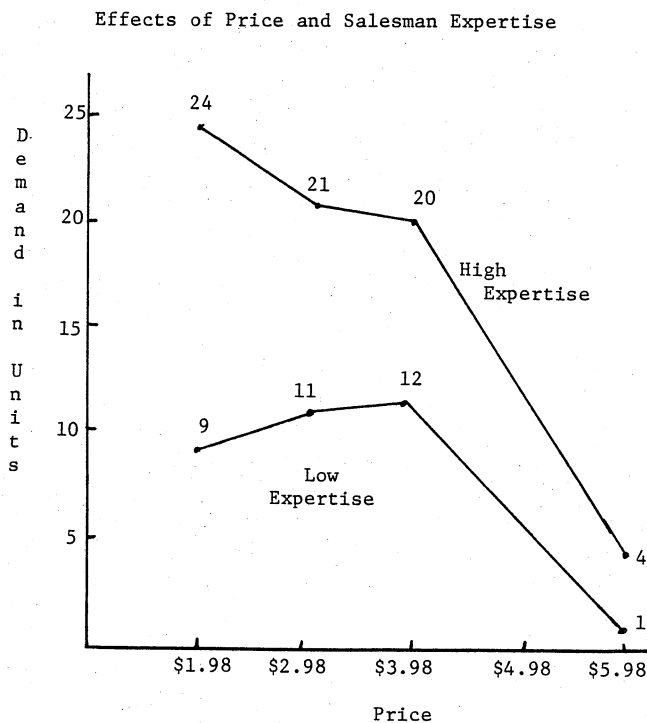
$$Z'_y = .582Z_E + .261Z_{(PrP)} - .268Z_{(EPr^2)}$$

$$Z'_y = .788Z_E + .260Z_{(PrP)} - .460Z_{(EPr)}$$

The main effects of price and pal were not statistically significant from an analysis of variance of the data using three prices. The price-by-pal ( $p < .001$ ) and expertise-by-price-squared ( $EPr$ ,  $p < .07$ ) interactions were statistically significant.

The interactions effects on purchases of price and expertise are apparent in Figure 5. For four prices

FIGURE 5



and high expertise the significant stepwise multiple regression equation is  $Z'_y = -.515Z_{Pr^2}$  and for four

prices and low expertise  $Z'_y = -1.480Z_{Pr^2} + 1.227Z_{Pr}$  a positive price slope occurs from \$1.98 to \$3.98 for the low expertise condition versus a negative price-squared slope for the high expertise condition.

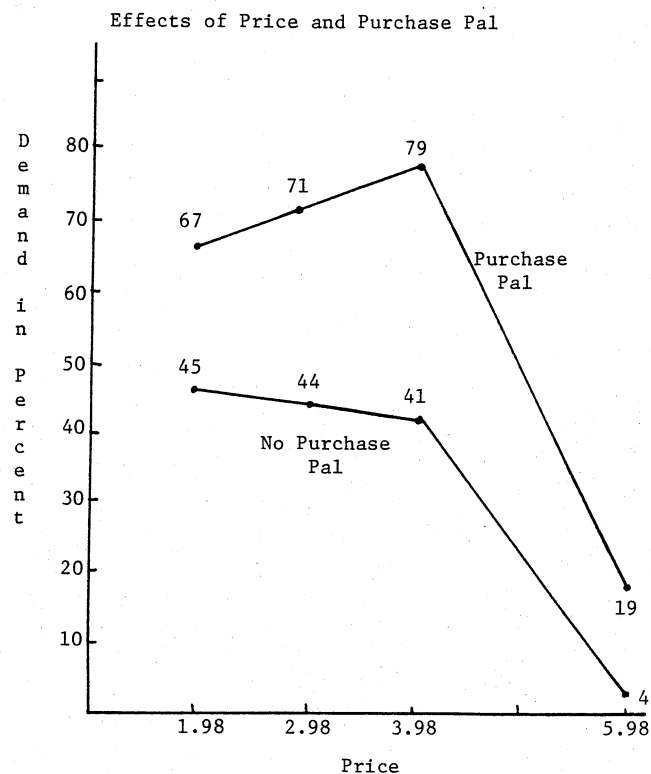
For the purchase pal conditions and four prices:

$$Z'_y = -.361Z_{Pr^2} \quad (\text{No Pal})$$

$$Z'_y = -.868Z_{Pr^2} + 1.525Z_{Pr} \quad (\text{Pal})$$

The quadratic relationship between price and demand for the four price treatments when purchase pals were present is shown in Figure 6. Both stepwise multiple regression functions for the absence or presence of purchase pals were statistically significant ( $p < .01$ ).

FIGURE 6



While some interaction terms were found to be significant,  $H_{4-7}$  were not supported by the analysis. For  $H_4$ , decreases in the likelihood of purchase produced by increases in prices were not greater, the lower the level of perceived expertise of the salesman. In fact, purchases increased with price increases from \$1.98 to \$3.98 for the low expertise condition.

Decreases in the likelihood of purchase produced by increases in prices were not greater when a purchase pal accompanied the customer compared with decreases when no purchase pal is present. In fact, purchases increased with price increases from \$1.98 to \$3.98 when a purchase pal was present.

## Conclusions

Customers had little opportunity to have prior knowledge concerning the product used in this study and they may be assumed to have engaged in Extensive Problem Solving decision making (Howard and Sheth, 1969), i.e., well-defined and structured choice criteria have not yet developed for the customers. Consequently, factors other than price, e.g., salesman expertise and advice of purchase pals, may affect purchase behavior to a greater extent, within a range of relevant prices.

Purchase pals may have a positive influence on customer purchasing behavior for some situations. Previous research (Life Insurance Agency Management Association, 1966) has found that when the wife is present during a sales presentation of life insurance, the likelihood of purchase by the husband is substantially increased (47% versus 32% in the study). Percent of purchases increased to 56% when the wife took an active part in the sales discussions. Thus, the presence of purchase pals in Extensive Problem Solving situations might be expected to produce greater likelihoods of positive sales outcomes. However, further research is necessary for different products, consumers, and decision making processes before any firm conclusions can be reached.

Price changes had no substantial effects on demand from \$1.98 to \$3.98, while demand was substantially decreased when the product was priced at \$5.98. This finding indicates that a broad range of acceptable prices may exist for consumers for some products, i.e., products for which consumers do not have prior price information.

More formal attention to significant interaction effects of marketing decision variables and other variables is needed in the study of consumer behavior. The study of such interaction effects will help pinpoint the conditions when one "common-sense" or intuitive belief will occur versus an opposite but equally plausible belief.

## Limitations

The findings presented in this study are limited to the customers participating in the experiment. Other forms and degrees of the expertise treatments could easily have been developed. Unfortunately, the discussions between the purchase pal and the customer, if any, were not recorded in the study. Also, Capon (1975) found a significant interaction effect between salesman and sales treatment which indicates that perception of and success of the salesman cannot be separated from the sales message (and vice versa). Thus, future studies of salesman-buyer interactions should include more than one salesperson and the resulting data should be analyzed for possible interaction effects between salespersons and sales messages.

## References

- Gerald D. Bell, "Self-Confidence and Persuasion in Car Buying," Journal of Marketing Research, 4 (February, 1967), 46-52.
- James R. Bettman, "Threshold Model of Attribute Satisfaction Decisions," Journal of Consumer Research, 1 (September, 1974), 30-35.
- Timothy C. Brock, "Communicator-Recipient Similarity and Decision Change," Journal of Personality and Social Psychology, 1 (June, 1965), 650-654.
- Noel Capon, "Persuasive Effects of Sales Messages Developed from Interaction Process Analysis," Journal of Applied Psychology, 60 (April, 1975), 238-244.
- Ronald C. Curhan, "The Effects of Merchandising and Temporary Promotional Activities on the Sales of Fresh Fruits and Vegetables in Supermarkets," Journal of Marketing Research, 11 (August, 1974), 282-294.
- Paul E. Green, "On the Analysis of Interactions in Marketing Research Data," Journal of Marketing Research, 10 (November, 1973), 410-420.
- John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior (New York: John Wiley, 1969).
- Paul Kau and Lowell Hill, "A Threshold Model of Purchasing Decisions," Journal of Marketing Research, 9 (August, 1972), 264-270.
- Herbert C. Kelman, "Processes of Opinion Change," Public Opinion Quarterly, 25 (Spring/1961), 57-78.
- Herbert C. Kelman and Alice H. Eagly, "Attitude Toward Communicator, Perception of Communication Content and Attitude Change," Journal of Personality and Social Psychology, 1 (January, 1965), 63-78.
- Philip Kotler, Marketing Management (Englewood Cliffs: Prentice Hall, 1972).
- Philip Kotler, Marketing Decision Making (New York: Holt, Rinehart and Winston, 1971).
- Philip Kotler, "Marketing Mix Decisions for New Products," Journal of Marketing Research, 1 (February, 1964), 43-49.
- Life Insurance Agency Management Association, The Family Financial Officer (Hartford: Life Insurance Agency Management Association, Report 5, 1966).
- John M. McCann, "Market Segment Response to the Marketing Decision Variables," Journal of Marketing Research, 11 (November, 1974), 399-412.
- Richard D. McKelvey and William Zavoina, "A Statistical Model for the Analysis of Ordinal Level Dependent Variables," Journal of Mathematical Sociology, 4 (in press, 1975).
- Arch G. Woodside and J. William Davenport, Jr., "The Effect of Salesman Similarity and Expertise on Consumer Purchasing Behavior," Journal of Marketing Research, 11 (May, 1974), 198-202.
- Arch G. Woodside and J. William Davenport, Jr., "The Effect of Price and Salesman Expertise on Customer Purchasing Behavior," Journal of Business (in press, 1975).
- Arch G. Woodside and J. Taylor Sims, "A Retail Experiment in Pricing a New Product," Journal of Retailing, 50 (Fall, 1974), 56-65.
- Arch G. Woodside, Robert E. Pitts, and Aaron Gewirtz, "Multiple Regression Analysis of Interaction Effects in Multidimensional Category Data," unpublished working paper, College of Business Administration, University of South Carolina, 1975.

Noel Capon and James M. Hulbert  
University of California, Los Angeles

Introduction

Interpersonal persuasion is a pervasive component of human behavior. However, its ubiquity has not been matched by corresponding research interest on the part of social psychologists. In the field of consumer-behavior, at both the individual and organizational level, interpersonal persuasion in the form of personal selling is an activity of great significance to the national economy. Yet, even here, the volume of research on personal selling is exceedingly small. (Silk & Davis, 1972; Capon, Holbrook and Hulbert, 1975).

A number of factors may account for this state of affairs. First, conceptually and methodologically, the study of such processes poses formidable challenges. Second, the relatively small number of researchers currently active in the area have been dispersed geographically, and have not possessed a vehicle to facilitate the exchange of experience and viewpoint. Third, some have viewed the subject as an unfashionable and even unpalatable research topic which lacks legitimacy.

The purpose of this session was to reduce these barriers. By having seven papers presented with a theme of Models and Methodologies for the Study of Interpersonal Interaction and Persuasion Processes, it was hoped that some of the conceptual and methodological issues could be explored and steps taken in the direction of increased understanding. In addition, both participants and audience might be stimulated by the papers and discussion to develop their own thinking and move into the area as researchers. Finally, the very existence of the session in this conference could help provide legitimacy for study in the area.

Models of the Interpersonal Persuasion Process

Five of the papers were concerned with the development of models and theories which attempted to describe and explain the interpersonal persuasion process.

The most all-embracing view was provided by Sheth, whose comprehensive model of buyer-seller interaction developed the concept of buyer-seller compatibility with respect to both content and style of communication. Employing a multiple interaction persuasion paradigm, Wilson developed a process model which employed the concepts of exchange theory (Homans, 1961) in a multi-attribute utility model format. The pervasiveness of multi-attribute theory in the consumer behavior literature was further reflected in the paper by Lutz and Kakkar, who, while using a situational paradigm, nevertheless cast their model in the form of the extended Fishbein model. (Fishbein & Ajzen, 1975).

Olshavsky advocated viewing the salesman-buyer interaction in terms of information processing theory (Newall & Simon, 1972). Perhaps reflecting his previous work

Noel Capon is assistant professor, Graduate School of Management, UCLA and James M. Hulbert is associate professor, Graduate School of Business, Columbia University. They wish to express their gratitude to the authors participating in the session, to Professors John Howard and Hal Kassarjian for their contribution as discussants and to those who generously gave of their time to review papers for the session.

with buyer-seller interaction data (Olshavsky, 1973), he argued for the development of process models based on individual interactions, an approach analogous to that of Bettman's work in individual consumer choice. (Bettman, 1970).

Finally, Holbrook & O'Shaughnessy, drawing upon the existing body of research on interpersonal persuasion, sought to model the influence process by using Kelman's influence typology (Kelman, 1961) and Etzioni's (1964) means to power.

The theoretical approaches were clearly quite different, although some convergence was evident. Thus most authors were concerned with the content or process of interaction, while the employment of the multi-attribute framework was a feature of some of the papers. Most papers drew heavily on concepts and constructs drawn from the existing literature, while Sheth, attempted a fresh approach at conceptualizing buyer-seller interaction.

Methodology

Just as various approaches to theory building were observed, so a number of methodological directions for gaining understanding of interpersonal persuasion were presented.

Of the seven papers, two, those of Sternthal, Scott and Dholakia, and Woodside and Pitts, reported empirical results, and methodological convergence was obtained through the use of the experimental paradigm. The utility of the experimental method for study of personal selling has been previously demonstrated (Capon, 1975; Farley and Swinth, 1967; Woodside and Davenport, 1974, and the widespread use of field settings is one of the most encouraging aspects of this work. Woodside and Pitts' study was in the classic tradition of communication research, price and communicator expertise being manipulated as independent variables and purchase observed as criterion. Sternthal, Scott and Dholakia's underlying theoretical perspective was quite different, and they reported the results of two studies which tested the predictions of self perception theory.

Holbrook and O'Shaughnessy argued for theory development before data collection, but approached the problem from an observational perspective via the recording of verbal content of interaction. They advocated the development of verbal content analytic schemes congruent with the researchers' theory of interpersonal influence, although they suggested that the coding of every verbal act, as in the Bales' (1950) system, might be too cumbersome and that more simplistic schemes related to the rules used by the actors in the buyer-seller interaction may be more practical.

Olshavsky, by contrast, suggested an inductive methodology in which data is collected via a two stage procedure. He was skeptical of the amount of information recovery possible with only an audio record and therefore suggested video, as well as audio, recording so as to retain a complete history, verbal and nonverbal, of the interaction (Hulbert and Capon, 1972). These recordings would then be played back to each participant so that a detailed retrospective analysis could be obtained.

From the individual analyses micro-models are developed of information processing by each participant during the interaction. The development of a series of such models would then allow for generalization and the development of macrotheories.

The other papers were concerned more with theory than methodology, though many of the ideas developed would seem amenable to testing by experimental modes. Such methods do of course have the considerable practical benefit of avoiding the labor-intensiveness associated with analysis of content.

#### Conclusions

Our review of the papers and of the session itself convinces us that the initial objectives were achieved. Equally evident, however, is the conclusion that the field is still at the very rudimentary stage of development. Definitive theory development is still awaited, and although the ideas presented in these papers suggested directions for future research, many questions remained unresolved.

One such question was raised by one of the discussants, who took the view that the considerable volume of persuasion research in the mass communication literature could be applied without serious alteration to the area of interpersonal persuasion. While few would argue that this literature could serve a very fruitful role in theory development, or that it could provide definite and immediate guidelines for experimental research, (viz. the papers by Woodside and Pitts; Sternthal, Scott and Dholakia), the general thrust of the papers and the session tended to view the interpersonal communications problem as more complex. The dynamic nature of interpersonal communication would suggest that more flexible persuasion strategies than are possible in mass communication might be evident, and that the kind of matching discussed by Sheth, and Holbrook and O'Shaughnessy may well be important. Indeed, this question is one that could well be approached empirically (Capon, 1975) and might well provide insight into persuasion processes in both mass and interpersonal contexts.

A second question was raised for those researchers who wish to focus their efforts on analysis of interaction data, for there are clearly some important ethical considerations to be resolved. Changes in both professional values and the law means that social scientists must, of necessity, be more circumspect in their field research. Appropriate means of reconciling subjects' and researchers' interests are an essential component in any design involving collection of interaction data.

It is our hope that this collection of papers and the interaction involved in their development, presentation and discussion will serve to stimulate further research effort. Interpersonal persuasion and communication research has for too long been the step-child of consumer behavior.

#### References

- Robert F. Bales, Interaction Process Analysis, Cambridge, Mass.: Addison-Wesley, 1950.
- James R. Bettman, "Information Processing Models of Consumer Behavior," Journal of Marketing Research, 7 (August, 1970), pp. 370-376.
- Noel Capon, "Persuasive Effects of Sales Messages Developed from Interaction Process Analysis," Journal of Applied Psychology, 60 (April, 1975), 238-244.
- Noel Capon, "The Relative Effectiveness of Interpersonal and Mass Communication: A Field Experiment," Working Paper, Graduate School of Management, UCLA, 1975.
- Noel Capon, Morris Holbrook and James Hulbert, "The Selling Process: A Review of Research," Working Paper, Graduate School of Business, Columbia University, 1975.
- Amitai Etzioni, Modern Organizations (Englewood Cliffs, N.J.: Prentice Hall, 1964).
- John U. Farley and Robert L. Swinth, "Effects of Choice and Sales Message on Customer-Salesman Interaction," Journal of Applied Psychology, 51 (April, 1967), 107-110.
- Martin Fishbein and Icek Ajzen, Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research (Reading, Mass.: Addison-Wesley, 1975).
- George C. Homans, Social Behavior: Its Elementary Forms (New York: Harcourt, Brace and World, 1961).
- James Hulbert and Noel Capon, "Interpersonal Communication in Marketing: An Overview," Journal of Marketing Research, 9 (February, 1972), 27-34.
- Herbert C. Kelman, "Processes of Opinion Change," Public Opinion Quarterly, 25 (Spring, 1961), 57-78.
- Allen Newell and Herbert A. Simon, Human Problem Solving (Englewood Cliffs, N.J.: Prentice Hall, 1972).
- Richard W. Olshavsky, "Customer-Salesman Interaction in Appliance Retailing," Journal of Marketing Research, 10 (May, 1973), 208-212.
- Alvin J. Silk and Harry L. Davis, "Interaction and Influence Processes in Personal Selling," Sloan Management Review, 13 (Winter, 1972), 59-76.
- Arch G. Woodside and J. William Davenport, Jr. "The Effect of Salesman Similarity and Expertise on Consumer Purchasing Behavior," Journal of Marketing Research, 11 (May, 1974), 198-202.

CONSUMER RESEARCH IN URBAN TRANSPORTATION:  
SOME METHODOLOGICAL ISSUES

Christopher H. Lovelock, Harvard University

Abstract

The nature of the transportation market has important methodological implications for consumer researchers. This paper points up opportunities for improving sample design, considers alternative methods of data collection, and discusses selection of measurement techniques. In the context of developing importance weights for multi-attribute models of modal choice, a review is made of priority evaluation and other methods.

Previous Research Traditions

Historically, transportation research has been dominated by engineers and economists. Engineering was the discipline which designed and built the physical facilities; it used analogies from the physical sciences to describe and evaluate traffic flows; and it employed often simplistic mathematical models to project the future demand for travel by alternative modes.

Economists were intrigued by the nature of the transportation industry, with its "perishable" product that could not be stored, its widely fluctuating demand levels over time, and its high fixed to variable cost ratios, as well as by the fact that the marginal costs of carrying an extra passenger under conditions of excess capacity tended to approach zero. These characteristics posed--and still pose--fascinating pricing problems for both passenger transit and public highway facilities.

Economic utility theories, meantime, provided the basis for development of econometric models which attempted to explain and predict consumers' choice of travel modes in terms of a very limited number of variables, notably price and travel time.

A New Marketing Orientation

In recent years, however, transportation studies have increasingly employed concepts and techniques drawn from the behavioral sciences.

In part, this represents recognition by transportation specialists of the insights to be gained from studying consumer behavior. It also reflects a growing interest by consumer researchers in applying their skills to new areas, outside such traditional fields as private sector marketing and political polling.

Finally, it represents a significant change in the whole transportation environment. Planners are no longer thinking in passive terms of projecting future travel demand patterns and then developing new facilities to meet forecast demand. Instead they are considering how demand patterns may be influenced and different modal choice behavior encouraged. This calls for an approach which considers the problem of moving people, with all their individual idiosyncrasies, rather than one of simply facilitating the movement of mindless vehicles. At the same time, a new breed of transit manager is emerging, who no longer speaks of riders as did one old timer, who was wont to refer to them as "seated freight."

With government encouragement, the transit industry is

adopting a new marketing orientation. Transit is now seen as a consumer product which must be sold against fierce competition. Efforts are therefore being made to tailor transit operations more closely to consumer needs and preferences, while monitoring and evaluation programs are being established to measure the effectiveness of particular managerial strategies in the transportation marketplace.<sup>1</sup>

Consumer Research Issues and Applications

As a result, research techniques developed in consumer marketing and political polling are finding increasing application in the urban transportation field. They are being used to determine not merely the size of the travel market but also ways in which it is segmented; to evaluate public attitudes towards both proposed and existing facilities; and to develop a better understanding of the factors underlying modal choice behavior and how consumers' decisions may be influenced.

It is encouraging to see marketing research tools and concepts being applied in the transportation field. During the past few years, a growing number of research studies have been conducted, often yielding significant findings. However, I believe that there is still plenty of room for further refinement of research procedures. Indeed, without such refinement we are unlikely to advance our understanding of modal choice behavior much beyond its present level.

In the balance of this paper, I would like to look briefly at some of the methodological issues facing urban transportation researchers. My objectives are, first, to point up some opportunities for improving sample design; second, to review the problem of selecting appropriate measurement techniques; and, third, to consider alternative methods of communicating with subjects.

These three issues are, of course, interrelated. The nature of the sample design may facilitate or impede certain forms of communication with members of the sample population; the characteristics of people in the sample frame may constrain the type of research instrument used and the measures employed; while the measures used may serve to determine--or be constrained by--the method chosen for communicating with respondents.

Sample Design

My first concern centers around, but is not confined to, issues relating to sample design. I sometimes wonder if many researchers who approach transportation from a consumer goods orientation truly understand the nature of the transportation "product."

Blumer (1969) has argued that methodology embraces the entire scientific quest and not merely some selected portion or aspect of that quest, such as data analysis. He emphasizes that "the entire act of scientific inquiry is oriented and shaped by the underlying picture of the empirical world that is used." My concern is that many

<sup>1</sup>See, for example, U.S. Department of Transportation (1975).

researchers do not have as clear and realistic a picture of the transportation world as is needed to conduct incisive research.

In evaluating modal choice decisions, there has been a tendency to conduct research and build models which see urban travel choices as lying simply between automobiles and public transportation. In practice, of course, there may also be other options such as walking or bicycling. Perhaps more serious is the practice of lumping all modes of public transportation together under one heading when comparing them with car travel (e.g. Paine et al., 1967; McMillan and Assael, 1968, 1969).

Many cities offer multiple-mode transit systems and it is most unrealistic to assume that consumers perceive different public modes such as buses, trains, rapid transit, streetcars and even ferries as broadly similar. My own research indicates that people see bus and train travel as differing sharply from one another on several important attributes (Lovelock, 1973). There is a risk that respondents may assign to the block term "public transportation" the characteristics of that mode which performs worst (or best) on any given attribute, particularly if questions are phrased in terms of asking them "how satisfied" they are with transit's performance on specific attributes.

Recently, I have come to the conclusion that it may also be dangerous to generalize within a given mode. To borrow a refrain from recent aspirin commercials, "all bus journeys are not alike." Many transportation studies distinguish between the nature of travel for different purposes, but they do not consider possible differences in the nature of the travel experience between, say, different routes.

#### Two Scenarios

Consider the following two scenarios. Mr. X and Ms. Y, who are quite closely matched demographically, both live in the same neighborhood, approximately the same distance from their respective bus stops. Ms. Y has a seat and enclosed shelter at her stop, whereas Mr. X has neither and is exposed to the elements.

Mr. X takes the #19 bus to his office. It is a notoriously unpunctual service. In its wisdom, the transit company habitually puts on an ageing vehicle with cracked windows, broken suspension and torn seats, driven by a bad-tempered operator who takes it out on passengers and vehicle alike. The bus is always packed and Mr. X has to stand the whole way. The route follows city streets up and down several steep hills, with many stops for congestion, traffic lights and embarking or disembarking passengers. The five mile journey takes Mr. X anywhere from 15-30 noisy, uncomfortable minutes, depending on the state of the traffic, the weather and the mood of the driver. Because of crowding and inadequate ventilation, the bus often smells like a zoo by the time Mr. X reaches his destination, usually feeling exhausted and either too hot or too cold.

Ms. Y also travels five miles to her job which, luckily for her, is in a suburban complex in the opposite direction. Buses on the #23 route run on time, are almost brand new and their drivers generally friendly and courteous. Running counter to the main commute direction, the bus is half empty and the route follows a smooth, limited-access highway with little traffic. The seats are comfortable, the heating and air-conditioning work, and it's a pleasant, 10-minute ride. By a quirk of the fare zone boundaries, Ms. Y's fare is 25 cents versus 40 cents for Mr. X.

In a recent survey, Mr. X gave bus travel a rotten rating on almost every attribute and indicated that he and his wife were thinking of buying a second car so he could drive to work. Ms. Y, by contrast, gave bus travel high marks all around and indicated that she planned to continue using it and was promoting it to all her friends.

#### Sample Design and Service Characteristics

Obviously, I have painted an extreme case in this fictional tale. But it does illustrate the possibility for wide variations in respondent attitudes towards what are effectively two very different products, although usually treated by researchers as one and the same thing. The nature of the "augmented product" varies widely in urban transportation and, like many services, is difficult to standardize. I believe that few other fields can offer the incautious researcher such opportunities for comparing apples with oranges, bananas with cucumbers, and coconuts with pumpkins.

What I am arguing for, therefore, is a much more incisive approach to sample design in urban transportation surveys. Researchers might usefully seek to stratify sample populations according to the observed conditions under which they generally travel. This calls for careful identification of subjects' most frequently made journeys, with special attention to the characteristics of routes which they use (or potentially might use) on a regular basis. Note that this applies to auto users as well as transit users. A car driver who regularly sees Mr. X's elderly #19 bus grinding up the hill is likely to form a poorer impression of bus travel than does the driver who sees the sleek new #23 speeding along the expressway.

Offhand, I do not know of any survey research which has attempted to be this specific. However, my research in the San Francisco Bay Area showed that respondents in different counties had significantly different perceptions of bus travel in their local areas, whereas there were no significant differences in their perceptions of car travel. These findings were consistent with my own more objective evaluations of the relative quality of service provided by the three different transit operations under review (Lovelock 1973).

#### Sample Design and Geographic Location

Surveys which seek to measure, say, public attitudes towards new tax proposals for financing transit should realistically sample the entire population of prospective taxpayers. However, surveys which seek to determine consumer satisfaction with specific attributes of public transportation, with a view to better understanding modal choice behavior, need to focus on geographically specific sample populations.

In a 1968 study, McMillan and Assael asked respondents, a high proportion of whom lived many miles from the nearest transit route, how satisfied they were with specific attributes of public transportation. This is hardly guaranteed to produce useful findings. It is an established fact that accessibility to a transit route of a traveler's origin and destination points is a significant factor in consumers' modal choice behavior (Lovelock, 1973, 1975). If we want realistic and meaningful evaluations of public transit modes against the private car, then a strong case can be made for confining our research to sample populations whose location provides a reasonable transit alternative.

Ideally, this calls for taking a map of the area where the survey is to be conducted, marking in transit routes, and then drawing contour lines around each stop or station to indicate accessibility in terms of travel

time to the stopping point. Studies suggest that most commuter rail patrons are located within a maximum of ten minutes' drive from the station, while the effective catchment area (or "transit envelope") for local bus service lies within 1000 feet or a two block radius of the bus stop (Bonsall, 1971; Lovelock, 1973; Metro Toronto Area Regional Transportation Study, 1969).

#### Excluding Unwanted Subjects

Depending on the objectives of the survey, it may also be necessary to narrow the sample frame on other than locational grounds.

Supposing we are studying the relationship between attitudes, beliefs and modal choice behavior for commute travel. Obviously, we want to exclude people who do not commute at all. This means removing retired people and those without regular jobs from the sample. We also need to exclude those whose jobs or physical condition clearly restrict them to a single mode, as well as those who do not have access to a car.

If we are using personal or telephone interviews, some preliminary questions may result in early termination of the interview on the grounds that a particular respondent falls outside the sample frame. Mail questionnaires can either be discarded after completion on the basis of responses to key questions, or subjects may be advised in the cover letter that they should only complete and return the survey if they meet certain criteria. Since the latter method may sensitize respondents or be misinterpreted by desired subjects, the former approach is recommended.

#### Selecting Measurement Techniques

The next issue I want to examine is that of selecting measurement techniques, particularly as these apply to measuring consumer attitudes.

During the past decade, attitudinal surveys have proliferated in transportation research. Stearns (1975) notes that "Attitude surveys have been widely used, frequently on an 'ad hoc' basis, for a variety of transportation planning and evaluation purposes. Their casual use has been due to a widespread need to measure motivational and evaluative factors and to estimate potential responses to alternative transportation services. These applications have been made despite a lack of a unified methodology integrating available and suitable techniques from several disciplines."

Unlike some areas of research in fields such as psychology or sociology, there are no well-tested "off-the-shelf" scales which researchers can use for measuring consumer attitudes in urban transportation. Development of common scales which could be used in various studies, conducted at different times and in different locations, might improve the generalizability of transportation research findings. While transportation studies have yet to reach this point, some researchers are beginning to replicate aspects of previous studies, which is a step in the right direction.

At a recent workshop, comprising participants drawn from transportation planning, transit system operation, market research and the social sciences (Stearns, 1975), three roles were outlined for attitudinal survey techniques in transportation planning and evaluation:

- (1) A marketing role which would measure transportation system and service preferences, market knowledge, and experience with the service;
- (2) An evaluative role to measure qualitative responses

to transportation system alterations; and

- (3) A planning role to measure local interests as an input to planning transportation systems.

#### Attitude Components<sup>2</sup>

When discussing attitudes, it is important to distinguish between their different components. There is wide acceptance of the concept that attitude structures have three component parts.

The cognitive or perceptual component represents a person's information about an object. Pieces of information can be broadly classified as either beliefs in the existence of an object or evaluative beliefs about an object. Techniques for measuring the former center on awareness measures, such as aided and unaided recall. Evaluative beliefs, meantime, provide information on the comparative judgements consumers make between alternative brands or products, notably as these relate to an individual's perceptions of specific attributes.

The affective or feeling component deals with a person's overall feelings of like or dislike for a situation, object, person or concept. By measuring consumer preferences, researchers may hope to obtain an understanding of the respondent's ideal product.

The conative component is concerned with an individual's intentions. Most research on attitudes has emphasized either their explanatory or predictive value, although the utility of the construct depends on achieving both. What has been termed the explanatory or structural approach deals with the cognitive or affective components, whereas predictive studies focus on the relationship between affect, intentions and overt behavior.

#### Multi-Attribute Attitude Models

Only recently have researchers come to appreciate the complexity of the transportation "product." There is now a recognition that consumers may choose between alternative modes on the basis of a variety of different attributes.

At last year's ACR meeting, I proposed a microanalytic model of the modal choice decision process (Lovelock, 1974). This saw consumers as trying to satisfy a particular travel need by first specifying the characteristics of the trip itself, then specifying the "ideal" modal attributes required for this trip, next evaluating the perceived characteristics of alternative modes against this "ideal" solution, and finally selecting that mode which provided the best perceived match.

Wilkie and Pessemier (1973) define a multi-attribute product as "a bundle of attributes leading to costs and benefits of differential desirability to individuals or segments of the market." Some of these attributes may be more important to consumers than others and in evaluating alternative products, consumers are often forced to make trade-offs between different attributes. It is essential that transportation planners and managers should not only be able to identify the attributes desired by consumers in modal choice decisions, but should also understand the degree of importance attached to each of these various attributes and whether or not different consumer segments set different priorities.

#### Identifying Salient Attributes

What attributes do travelers consider important in selecting a mode of transportation? A number of studies

<sup>2</sup>This discussion is drawn primarily from Day (1973).



have researched this issue and have come up with broadly similar conclusions.

Solomon, Solomon and Silién (1968) reviewed five different research studies conducted in various locations in the United States between 1962 and 1967. Depending on the study, subjects were asked either to rank specified attributes in order of importance, to list attributes they considered important, or to rate specified attributes on Likert scales. Solomon *et al.* noted that, despite significant differences in methodology and often widely varying characteristics in the sample populations, the ranking of modal choice criteria was unusually consistent across all five surveys. They summarized these criteria, in order of importance, as safety, reliability, time savings, cost, convenience and comfort.

Among the more sophisticated of these five studies was that of Paine *et al.* (1967) who surveyed 1,021 respondents in Baltimore and Philadelphia. Based upon a thorough review of previous research findings they developed an attitude instrument which would measure consumer attitudes along 33 variables believed to relate to modal choice decisions. Respondents were first asked to rate the importance of each item on a 7-point scale. Items consisted of descriptive phrases and were scaled from "Not at All Important" to "Of Greatest Importance." Factor analysis was subsequently employed to reduce these 33 items to a more parsimonious set of variables. From this analysis the researchers were able to suggest the basic attributes of a generalized ideal system.

McMillan and Assael (1968, 1969), conducted a national survey of 2,500 respondents replicating several aspects of the earlier Paine study. They used a very similar format, with 7-point scales and self-administered questionnaires, but only 15 descriptive phrases instead of 33.

A somewhat different approach to measuring importance of transportation attributes was followed by Golob *et al.* (1972). As part of their research, they conducted a home interview survey in the suburban city of Warren, Michigan. Some 800 respondents completed a questionnaire using the method of paired comparisons to evaluate 32 transportation system characteristics.

The results of this study, which was undertaken after Solomon *et al.*'s review, yielded very similar findings to previous studies, with the results emphasizing the attributes of punctuality/reliability, comfort, convenience, travel time and cost (in approximately that order). It may be noted that Golob *et al.* excluded attributes relating to safety from their study, on the grounds that safety of the system was not realistically subject to trade-off in transit system design. This raises the issue for researchers of deciding which characteristics considered "important" by respondents are actually salient attributes that should be included in a multi-attribute model of modal choice.

Alpert (1971) emphasizes that not all attributes rated as important by consumers are necessarily determinant, in terms of differentiating one product (or mode) from another. For instance, safety may well be taken for granted by many travelers when selecting among alternative modes of urban transportation. Alpert suggests three broad approaches for identifying determinant attributes, namely direct questioning, indirect questioning (including motivation research and covariate analysis), and observation and experimentation. His own research (on attitudes towards moderately priced pens) indicated the most efficient approach to be direct dual questioning, where respondents are asked to rate product attributes in terms of (1) how important

each is thought to be in determining choice, and (2) how much difference is perceived among competing products in terms of each attribute.

Perhaps we need to review this issue of salience more carefully as it relates to safety, since consumers may be concerned not merely with issues relating to safety of the vehicle on the track or highway but also with personal safety from the standpoint of being mugged at the bus stop.

We also need to consider the extent to which the relative importance of different attributes may vary according to the nature of the trip being made and the personal characteristics of the traveler. Paine *et al.* and McMillan and Assael found some differences between requirements for work and non-work trips. Golob *et al.* (1972) looked at differences between demographic groups and found some variations in the rank order of attributes for single people under 20, for the elderly and for low income groups, but not for any of the other segments they broke out. A later study by Golob, Dobson and Sheth (1974) showed limited differences according to respondents' sex and age; low income respondents (under \$6,000) evidenced different concerns from other income groups, and there were also variations between central city residents and suburbanites. However, to my knowledge, no in-depth study has yet attempted to examine variations in importance ratings with reference jointly to both trip purpose and traveler characteristics. I believe this constitutes an important research opportunity.

One issue which should be raised here concerns the most appropriate technique to employ when measuring importance. While experience shows a variety of different methodologies to have generated similar rankings, a purely ordinal scale does not yield the same insights as interval scale data.

The advantage of the method of pair comparisons employed by Golob *et al.* (1972) is that it forces respondents to compare each item separately against all other items, whereas Likert or semantic differential scales do not force such choices and allow for ties in ratings. However, the problem with this first method is that the number of pair comparisons increases at an exponential rate as the number of stimuli ( $n$ ) increases, the formula being  $n(n-1)/2$ . Although relatively straightforward, the process can quickly become time consuming and boring, leading to respondent fatigue.

When measuring consumer preferences for 32 different modal characteristics, Golob *et al.* were able to reduce the number of paired comparisons required from a single matrix of 496 to nine smaller matrices, each related to a specific group of characteristics and totalling only 168 choices. To provide a common basis for measuring the relative importance of all characteristics, several of them were included in more than one group. Nevertheless, correspondence with the researchers revealed that a 30-page questionnaire was required for these 168 paired choices and that respondents needed between 30 and 75 minutes to complete the questionnaire, with an interviewer in attendance to monitor the process.

It is evident that the paired comparisons method quickly becomes unwieldy for rating more than a limited number of attributes, and may need closer supervision than Likert or semantic differential scales. As a result, it appears to be of limited usefulness in large-scale surveys where the researcher has a variety of different questions to ask respondents and/or is operating under budget constraints which make supervised completion of questionnaires at home an unrealistic alternative. On the grounds of economy, speed and simplicity, I would therefore argue for use of Likert and semantic differen-

tial scales in transportation attitude studies.

#### Use of Constant Sum Scales

An alternative approach to measuring salience which also forces trade-off variations among attributes is the constant sum method, in which subjects are required to allocate a fixed number of points between each of the specified attributes. While I am not aware of any published transportation research employing this technique in the United States, it has been used for transportation and related studies in Britain (Hoinville, 1973; Hatch and Flack, 1974). I consider it a sufficiently interesting methodology to merit more detailed discussion.

Hoinville argues that the main limitation of most attitude measures is that respondents are not forced, as they are in a behavioral situation, to trade off some preferences against others. He asks, "Can people behave in a serious and responsible way when removed from the pressure of a real behavioral decision?"

The problem, as Hoinville sees it, is to develop a measuring device which maintains simplicity for respondents and yet approaches the complex reality of multi-choice situations. He developed some interesting priority evaluation techniques for his own research into community preferences concerning such environmental variables as traffic noise, parking availability, pollution levels, etc. The range of choices offered respondents was illustrated pictorially on a board containing 30 small pictures. Different drawings represented 3 separate standards for each of the 10 environmental variables, and were ranked from "poor" on the left to "good" on the right. These drawings had been pretested for comprehension and rank ordering by traditional interviewing methods. Although some of the concepts were admittedly difficult to depict, it could be argued that pictures conveyed more than could just a brief verbal description.

Next, a price tag was attached to each standard, representing the crude supply costs of arriving at each standard. Respondents were then given fifteen pegs, like radio jacks, each worth £100 and told that they must select one standard for each variable. The baseline (poor) standard was free; better standards could be "purchased" by inserting the appropriate number of pegs.

Respondents could easily assess their present position at any time, since existing choices on the board were illuminated electrically. At the outset, all baseline choices were illuminated. As the respondent "purchased" a more desirable position by inserting pegs, so the light in the old position would be extinguished and the new one lit up. Respondents were allowed to modify their choices after studying the outcome or "pay off" resulting from these choices. The advantages of this approach included ease of use for respondents, flexibility, avoidance of excessive mental arithmetic and a very wide range of choice combinations (the 10 x 3 matrix on the board provided for more than 3,700 possible combinations from which to choose). By changing the "costs" attached to the different alternatives, it was possible to see how respondents altered their positions on the board to generate a different set of trade offs.

Hatch and Flack used priority evaluation techniques (but without Hoinville's sophisticated electric board) to assess consumer preferences for alternative improvements in the London Underground. While they feel that the approach provides useful data on consumer priorities, they also note some general problems. In

particular, they emphasize that successful use of priority research techniques rests on an assumption that consumers actually think in terms of priorities and alternatives, whether or not financial and other constraints are placed on preferences. They argue that this may not be the case and believe there is a danger that "introducing alternatives and priorities to the sample being studied educates and conditions it so that it is no longer typical of the universe it was selected to represent."

Obviously, the priority evaluation techniques developed from constant sum scales are not without their disadvantages. Except for very simple scales, they tend to require the presence of an interviewer or supervisor, with accompanying increases in costs. Nevertheless, I believe that this methodology merits further consideration for transportation research purposes on this side of the Atlantic.

#### Measuring Consumer Awareness

As noted earlier, the cognitive or perceptual component of attitudes represents a person's information about an object. It can be divided into beliefs in the existence of an object (i.e., awareness) and evaluative judgements about the object.

It is self evident that if a person does not know that a product (such as transit service) exists, he or she can hardly be expected to use it except under impulse purchase conditions. However, it can be argued that mere knowledge of existence is scarcely sufficient in most situations, and that consumers need to have information about certain key attributes of a product before they can make comparative judgements against competing products.

A major shortcoming of economic theories of consumer behavior, and the models based upon these, is that they cannot explicitly handle the problem of imperfect information (Ratchford, 1975). It may very well be that a partial explanation for modal choice behavior patterns may be found in individuals' lack of knowledge of the availability and characteristics of alternative modes.

In the case of public transportation, prospective users may need to know about the routing of services (and, in particular, the location of stops in relation to their journey origins and destinations); what time the transit vehicle leaves and when it arrives; how much the journey costs and how the fare is to be paid.

Unlike attributes such as comfort, which are highly subjective and complex to measure (Nicolaidis, 1975), routing, scheduling and costs characteristics can be quantified in unambiguous terms. Consequently, respondents can be surveyed to determine how knowledgeable they are about the existence and specifics of transit services in relation to these attributes and their answers graded by degree of accuracy.

If the location of respondents is known (and this speaks to issues of both sample design and delivery procedures), then each can be asked such specific questions as (1) How far is your home from the nearest bus stop? (2) Is there a train service from your neighborhood to downtown? (3) How long does the train journey take on this route during commute hours? (4) How frequently does the bus operate on this route on weekday mornings outside commute hours? (5) How much is the bus fare from your home to downtown? (6) What is the name of the transit company providing the service? and so forth.

Methodological problems for the researcher center around formulating questions which can easily be validated for each respondent (which requires identifiable rather than

anonymous responses); are equally appropriate for all respondents in the sample frame (which may cover a wide geographic area); can be answered unambiguously by respondents without undue prompting; and are simple to code for subsequent analysis.

The problem of developing such measures is compounded in situations where transit services are not scheduled consistently on regular headways (e.g., every 30 minutes), respondents can choose between express and local services on the same routes, there are alternative transit routes between the same two points, and there are a variety of different fair options. However, it should not be beyond the wit of the ingenious researcher to circumvent such problems by careful attention to both sample design and questionnaire wording.

As an input to modal choice models, it may be useful to develop an awareness index for each consumer. For such a purpose, the researcher may wish to weight certain awareness items more heavily than others and also, perhaps, develop a scoring procedure which gives partial credit for responses within defined limits of the correct answer.

In my own research (Lovelock, 1973), I have used a simple additive awareness index, whereby respondents were given one point for each answer which was correct within certain limits (e.g., a fare within 10 cents of the actual fare), but nothing for wrong answers, and the totals then summed for each individual. While this measure discriminated quite well between users and non-users of transit (as one would expect), I believe that more sophisticated scoring and weighting procedures might improve the explanatory and predictive power of the awareness variable.

#### Problems With Halo Effects

Attempts to measure evaluative beliefs about specific attributes of an object may be confounded by halo effects. Several researchers have drawn attention to the problems associated with such effects in multi-attribute attitude models (Wilkie and Pessemier, 1973; Beckwith and Lehmann, 1975). The essence of the problem is that individuals who favor a particular product tend to rate it highly on all desirable attributes, while individuals who dislike it have a tendency to give it a low rating on each of the attributes listed. Wilkie and Pessemier note that "halo effects have long been recognized in personality and psychological testing as potential suppressors of importance variation. The presence of halo effects in the marketing model . . . will confound investigations as to the dimensionality of attitude structure and impair diagnostic analyses of brand strengths and weaknesses."

In a transportation context, the result of halo effects is most likely to be reflected in superior ratings for car travel (generally the preferred mode) across all attributes and inferior ratings across the board for travel by transit modes. What can be done to counter this problem?

Wilkie and Pessemier's suggestions include adding warm-up instructions to respondents which discourage yeasaying; and rating all products (modes) within attributes (rather than rating each mode separately on all attributes). Beckwith and Lehmann recommend that care be taken to include only relevant attributes, since respondents' statements about less important attributes may be determined almost totally by the halo effect. They also urge that researchers select attitude measures which are as objective as possible. Amongst other things, this argues against use of terminology such as "how satisfied are you?" in favor of a more neutral

"please rate performance." It is perhaps noteworthy that both Paine *et al.* and McMillan and Assael employed satisfaction measures in their respective studies, which showed respondents to rate car travel more favorably than transit on every single attribute. By contrast, Lovelock (1973) used semantic differential scales and found train and bus travel to be rated significantly better than car travel on the characteristic of safety.

#### Communicating With Subjects

In this section, I want to look at the problem of selecting the most appropriate means of communicating with subjects for transportation studies.

It is important to emphasize that every research program poses somewhat different needs. While the objective function in each instance should (hopefully!) be to maximize the quality and usefulness of the research findings, the constraints involved for any given project are likely to vary widely. Non-response bias may have more serious implications in some studies than in others; cost may or may not be a key consideration; and the importance of precise sample design varies with the characteristics of the universe under study. It can never be claimed that one methodology is always "superior" to another, only that one may be more suitable under certain, specific conditions.

As mentioned earlier, the choice of procedures for obtaining information from respondents is closely related to issues of sample design and selection of measurement techniques.

In evaluating alternative methodologies, the researcher must recognize the interrelationships between such decision variables as size, characteristics and location of size, degree of detail desired in the responses, questionnaire length, response rate required, size of total budget, anticipated cost per usable response, and survey topic. Changes in any one of these variables may impact upon others. Thus, a longer questionnaire should provide more detailed results but may be more costly to produce and mail; under some circumstances it may also result in a lower response rate, thereby generating not only higher total costs but also significantly higher costs per usable response. Attempts to increase response rates by monetary incentives, careful follow-up procedures or other techniques tend to raise total costs but the average cost per response received may not raise proportionately.

#### Choice of Locations for Surveys

The nature of the transportation market provides some interesting opportunities for reaching potential and actual users, in that the sample population is typically defined by such readily identifiable characteristics as usage behavior or origin and destination location. The researcher's problem is to determine the most appropriate means of reaching subjects in these sample populations. Opportunities exist for surveying respondents at home, at destination locations, or even en route as they travel between the two.

Let's examine each of these alternative locations in turn, and consider their implications for sample design, selection of measurement techniques, and general nature of the research instrument.

If we can reach people while they are actually traveling, we are immediately in a position to define our sample as people who behave in a specific way of interest to us (i.e., traveling that day at a specific hour, by a particular mode on a given route). Depending on the nature

of the mode and the travel conditions at the time, it may be possible to conduct on-the-spot personal interviews, ask travelers to complete a questionnaire for immediate or later return, or take individual names, addresses and/or phone numbers to allow for subsequent mail, phone or personal interviews. An interesting variation on the latter approach, occasionally used in highway traffic surveys, is to photograph cars' rear number plates, obtain names and addresses of their owners from state motor registry files, and then mail questionnaires to these people.

Most of us are probably familiar with cordon traffic surveys, where a proportion of cars passing a particular point are stopped and brief questions asked of their occupants. Similar surveys are sometimes conducted at transit stops and stations or on board moving vehicles. However, there are obvious limitations here on the type of instrument employed. Congestion on highways restricts the feasibility of traffic surveys and passenger congestion often makes it impracticable on board transit vehicles. Even where contact can be made, the amount of time which is available may be too short for all but a short, simple survey which cannot get into detailed attitudinal issues. Usually such surveys are designed to obtain descriptive information about travelers, such as their journey origins and destinations, trip purposes, key demographic characteristics and, perhaps, frequency of travel on that route by that mode.

On longer journeys by public transportation (such as trips on express buses, ferries and trains, or airline flights), the situation may be different. In these situations, the passenger is effectively a "captive," more likely to have a seat, less subject to other distractions and perhaps even responsive to the opportunity to pass the time by completing a written questionnaire.

The appropriateness of the second alternative, reaching people at their destination (which for our purposes may be considered the location reached on the outbound journey from the traveler's home residence), depends largely on the nature of the destination. There are obvious differences between interviewing respondents (a) as they arrive at a bus or rail terminus; (b) arrive at a bus stop; (c) leave their car at a parking lot; (d) start a shopping expedition at a particular location; (e) arrive at a theatre or sports stadium; or (f) arrive at work. Realistically, the first five situations probably allow for no more than a brief survey (if that), or for collecting names and addresses for subsequent follow-up.

Surveys at work locations have more potential, provided cooperation can be obtained from the employer in question. Possibilities include completion of either personal interviews or self-completion questionnaires on company time; or distribution of questionnaires by the employer for completion by employees during their own time, with return either through the mails or via the employer. However, employers may be unwilling to provide the researcher directly with a list of employee names and addresses for independent follow-up.

Because of the time constraints and other distractions inherent in en route and destination surveys, it is reasonable to anticipate that when conducting detailed attitudinal surveys, subjects' homes will be the preferred choice of the transportation researcher.

#### Evaluating Alternative Interviewing Media

Three basic media of communication are generally available to the researcher wishing to conduct interviews with subjects at their home residences, namely personal

interview, telephone interview and unsupervised completion of a questionnaire by the respondent. Variations on these three basic approaches include: the use of advance postcards or phone calls to inform people that a mail or personal survey will soon be conducted; a choice between using the mails or personal drop-off and pick-up for delivery and collection of questionnaires; and including a self-completion questionnaire as part of a personal interview session.

The major criteria usually employed in selecting between these alternatives include cost per completed response, the length and nature of the questions to be posed, the format in which responses are desired (e.g. open-ended versus forced choice), the desire to avoid various types of bias, the characteristics of the sample, and the time frame available for completion of the research. In this section of the paper, I propose to briefly review the appropriateness of each of the three major alternatives from the standpoint of attitudinal research within a transportation context.

#### Telephone Surveys

Telephone surveys are attractive to many survey researchers. They are normally less costly and more expeditious than face-to-face interviews, while still providing for interaction between interviewers and respondents. They generally achieve higher rates of response than either mail or personal contact and can be validated by monitoring or callback (Glasser and Metzger, 1972). However, I believe that telephone interviews may present problems which make them a less-than-satisfactory means of gathering detailed attitudinal data for transportation surveys.

A key problem relates to the risk of sample bias resulting from households without telephones and households with unlisted telephone numbers. Glasser and Metzger (1975) indicate that nationally, 19.2% of household subscribers had unlisted numbers in March, 1974. Potentially of serious concern to transportation researchers is the incidence of non-listings among specific demographic and geographic groups. It is particularly high in the West census region (26.1%); in counties located in the five largest metropolitan areas (29.0%) and the next 21 largest metro areas (23.5%); in households with incomes between \$5,000 - \$9,000 (25.0%); among non-white subscribers (31.7%); and among males aged 18-34 (24.5%) and females aged 18-34 (27.0%).

Since transportation research tends to focus on large metropolitan areas, the high incidence of non-listings in such areas should be cause for serious concern among those who generate their phone or address samples from telephone directories. And since public transportation tends to attract a higher proportion of people in lower income groups and from non-white ethnic groups, there is also a significant risk of bias from that standpoint. By contrast, those in the lowest income brackets (under \$5,000) have a relatively low incidence of non-listing (17.6%) as do those earning \$15,000 and over (17.3%), men aged 50 and over (8.1%) and women aged 50 and over (11.8%).

Glasser and Metzger (1972) suggest random digit dialing as a method of telephone sampling which avoids the non-listing problem. However, as emphasized earlier, transportation study samples should be geographically specific when studying modal choice behavior, since proximity of access to a transit route may be a key determinant. Unfortunately, in most locations the preliminary two or three digits of a number may delineate telephone subscriptions within quite large neighborhoods or even entire cities and are thus unsuitable for pinpointing sample locations with sufficient precision.

A second, and equally important problem with this medium is that the amount of information which may be obtained over the phone is typically limited. Although Payne (1974) has indicated some subtle ways of overcoming the telephone's shortcomings, it remains difficult to administer more than a limited number of questions employing psychological scaling methods such as paired comparisons, Likert scales or semantic differential scales through this medium.

#### Personal Interviews

Personal interviews, by contrast, are an appropriate and frequently used way of administering detailed attitudinal surveys, providing skilled personnel are available to minimize interviewer-induced response bias. However, they can prove extremely expensive to implement. Figures of \$20 to \$50 per completion are common where such interviews are conducted on an individualized basis. Where large sample sizes are desired for statistical purposes, this method of obtaining information may quickly become infeasible for all but the most generously funded projects, unless some way can be found to bring respondents together in a central location.

There are many works detailing the advantages and disadvantages of personal interviewing in survey research.<sup>3</sup> So I will not reiterate these here, except as in they relate specifically to transportation studies. Selection of households in personal interviews conducted for this purpose can be made on the basis of housing unit location relative to transit routes, major highways and distance from major employment centers, in consultation with demographic census data (which is available on a block by block basis). Additionally, the interviewer may be able to observe certain characteristics, such as physical condition of respondents, model and condition of household automobiles, etc., which might be related to modal choice behavior.

One problem with personal interviews nowadays, however, is concern for the safety of the interviewer, especially in neighborhoods with an above average crime rate. Additionally, many people are increasingly unwilling to open their doors to a stranger. These two factors may make it particularly difficult to interview respondents in the lower income brackets.

Never the less, certain techniques, such as prompted recall, paired comparisons and the priority evaluation methods employed by Hoinville (1973) and Hatch and Flack (1974) effectively require the presence of a personal interviewer. Direct questioning by an interviewer also has advantages when measuring consumer awareness since it does not permit time for extensive reflection by the subject nor provide an opportunity for cheating.

#### Self-Completion Questionnaires

If the telephone is deemed inappropriate and personal interviews considered too expensive, then the researcher is forced to use questionnaires designed for self completion by respondents. Again, I will refrain from discussing the general pros and cons of mail questionnaires, since these are treated in detail elsewhere in a variety of different articles and reference works.<sup>3</sup> Suffice to say that such questionnaires have proven to be an effective medium for often quite extensive attitudinal surveys on a wide variety of topics.

As in personal interviews, the sample may be defined with reference to transit route and stop locations, in conjunction with census data. Questionnaires may then be mailed directly to addresses on appropriate streets.

<sup>3</sup>Ferber (1974) is a good sourcebook.

An alternative delivery procedure calls for personal drop-off and pick up of questionnaires by lightly trained survey takers, following preassigned routes and visiting a given proportion of household units on that route. If desired, the number of questionnaires distributed to each household can be varied, to reflect the number of eligible subjects resident therein.

This approach also allows for selective replacement of subjects falling outside the sample frame (e.g., non-English speaking persons). Survey takers may be required to keep log sheets, allowing for identification of reasons for non-response and thus a better understanding of the nature of non-response bias than is possible in mail surveys. I have used this drop-off and pick up procedure successfully myself and it is described in detail in Lovelock (1973).

#### Conclusion

This paper has reviewed a number of methodological issues which I consider particularly significant for transportation researchers. I have deliberately chosen to focus on methods relating to research design and data collection rather than data analysis, since I believe that it is in the former categories that improvements are most needed and will yield the most significant advances. The analytical techniques employed in transportation research are, in many instances, highly sophisticated. However, the usefulness of any resulting findings is conditioned by the quality of data collected in the first instance, and I perceive many opportunities for improvement in sample design and measurement techniques. Researchers also need to exercise care in their choice of methods for communicating with subjects; in particular, they should recognize that telephone interviews may be inappropriate for certain types of transportation research.

By way of conclusion, I should like to make a plea for more detailed reporting in transportation studies of what I consider to be essential background information. Since the value of research findings is so closely tied to the nature of the methodologies employed, readers have a right to expect reports to provide information on such basics as the characteristics of the sample (as well as just its size), how subjects were selected, and the percentage response rate; it is also desirable to include an identification and evaluation of any possible biases. This information can generally be provided quite concisely but some significant background details are missing all too often.

One explanation for this situation may perhaps be found in the fact that in transportation studies, as in other fields, the tasks of data collection and data analysis are often entrusted to two different researchers. However, if this paper achieves no more than to lead the analysts to ask more pointed questions about where their data came from and how it was collected, it will have served a useful purpose.

#### References

- Alpert, Mark I., "Identification of Determinant Attributes: A Comparison of Methods," Journal of Marketing Research, 8 (May 1974), 184-191.
- Beckwith, Neil E. and Donald R. Lehmann, "The Importance of Halo Effects in Multi-Attribute Attitude Models," Journal of Marketing Research, 12 (August, 1975) 265-275.
- Blumer, Herbert, Symbolic Interactionism: Perspective and Method (Englewood Cliffs, N. J.: Prentice-Hall 1939).

- Ronsall, J.A., Dial-A-Bus: The Bay Ridges Experiment (Toronto, Ont.: Ontario Dept. of Transportation and Communications, 1971).
- Day, George S., "Theories of Attitude Structure and Change," In S. Ward and T.S. Robertson (eds.), Consumer Behavior: Theoretical Sources, (Englewood Cliffs, N.J.: Prentice-Hall, 1973).
- Ferber, Robert, Handbook of Marketing Research, (New York: McGraw-Hill, 1974).
- Glasser, Gerald J., and Dale D. Metzger, "Random Digit Dialing as a Method of Telephone Sampling," Journal of Marketing Research, 9 (February, 1972), 59-64.
- Glasser, Gerald J. and Gale D. Metzger, "National Estimates of Nonlisted Telephone Households and Their Characteristics," Journal of Marketing Research, 12 (August, 1975), 359-361.
- Golob, Thomas F., Eugene T. Canty, Richard L. Gustafson, and Joseph E. Vitt, "An Analysis of Consumer Preferences for a Public Transportation System," Transportation Research, 6 (March, 1972), 81-102.
- Golob, Thomas F., Ricardo Dobson, Jagdish N. Sheth, "Perceived Attribute Importance in Public and Private Transportation," (Warren, Mich.: General Motors Research Publication GMR-1465, July, 1974).
- Hatch, Donald and Mervyn Flack, "Priority Research Techniques Applied to Public Transport Investment," Journal of the Market Research Society, 16:1 (1974).
- Hoinville, G., "Evaluating Community Preferences," Journal of the Market Research Society, 15:1 (1973), 1-23.
- Lovelock, Christopher H., Consumer Oriented Approaches to Marketing Urban Transit, Ph.D. dissertation, Stanford University (Springfield, VA.: National Technical Information Service, #PB. 220-781, 1973).
- Lovelock, Christopher H., "Researching and Modeling Consumer Choice Behavior in Urban Transportation," in M.J. Schlinger (ed.), Advances in Consumer Research, Vol. II, 1974 Proceedings, Association for Consumer Research.
- Lovelock, Christopher H., "A Market Segmentation Approach to Transit Planning, Modeling and Management," Proceedings, Transportation Research Forum, 1975.
- McMillan, Robert K. and Henry Assael, National Survey of Transportation Attitudes and Behavior (Washington, D.C.: National Cooperative Highway Research Program Reports #49, 1968; #82, 1969).
- Metropolitan Toronto and Region Transportation Study (MTARTS), GO Transit Commuter Rail Project, Special Reports #4A and #9 (Toronto, Ont., 1968).
- Nicolaidis, Gregory C., "Quantification of the Comfort Variable," Transportation Research, 9 (1975), 55-66.
- Paine, Frank T., et al., Consumer Conceived Attributes of Transportation: An Attitude Study (Springfield, Va.: National Technical Information Service, #PB 176-485, 1967).
- Payne, Stanley L., "Data Collection Methods: Telephone Surveys," in R. Ferber (ed.) Handbook of Marketing Research (New York: McGraw Hill, 1974).
- Ratchford, Brian T., "The New Economic Theory of Consumer Behavior: An Interpretive Essay," Journal of Consumer Research, 2 (Sept., 1975), 65-75.
- Solomon, Kathleen M., Richard J. Solomon and Joseph S. Silien, Passenger Psychological Dynamics, Sources of Information on Urban Transportation, Report #3 (New York: American Society of Civil Engineers, 1968).
- Stearns, Mary Donahue, Proceedings: TSC Workshop on Attitudinal Studies for Transportation Planning and Evaluation (Cambridge, Mass.: Transportation Systems Center, 1975 [forthcoming]).
- United States Department of Transportation, Proceedings, National Transit Marketing Conference (Washington, D.C.: Office of Transit Management, 1975).
- Wilkie, William L. and Edgar A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10 (November, 1973), 428-441.

A BEHAVIORAL TRAVEL DEMAND MODEL INCORPORATING CHOICE CONSTRAINTS

Wilfred W. Recker,<sup>1</sup> State University of New York at Buffalo  
 Thomas F. Golob, Research Laboratories, General Motors Corporation

Abstract

Market segmentation techniques are used to capture the effects of availability constraints on urban residents' choice of automobile or bus modes of transport for their journey to work. Logit probabilistic choice models are then estimated for each market segment. The explanatory variables in these models are home-interview survey respondents' attitudes toward the two modal alternatives expressed in terms of their satisfactions with a series of descriptive attributes.

Introduction

Previous urban transportation research studies have shown that attitudes can be used as effective descriptors of consumers' travel preferences (Golob and Dobson, 1974). Considerably less success has been achieved in using attitudes as descriptors and predictors of actual choice of mode or destination of travel (Hartgen, 1974). It is hypothesized herein that this disparity in attitudinal descriptive power is principally attributable to the previous exclusion of intervening constraints affecting the realization of preferences in actual choice behavior. Such constraints in the case of modal choice decisions are typically associated with supply-side characteristics of the transportation systems defining the modal choice alternatives.

While attitudes toward these supply-side characteristics can be measured, it is in many cases inappropriate to consider that these attitudes enter the choice decision process in a compensatory manner (i.e., as additive components). As the most general case, it can be expected that individuals operating under different sets of supply-side constraints will both view their choice alternatives differently and also possess different relationships between their choices and their attitudes toward the choice alternatives.

To assess the impact of such supply-side constraints on individuals' choice of travel mode, choice models were developed which explicitly incorporate constraints in a non-compensatory manner. These models are then interpreted in terms of forecasting usefulness and are contrasted to a choice model not incorporating constraints. The explanatory variables in all of the choice models are travelers' attitudes toward their modal choice alternatives on a comprehensive set of descriptive attributes. The models are probabilistic, and are of the strict-utility genre.

The data was obtained from the Regional Municipality of Ottawa-Carleton, Ontario, Canada and was collected through a home-interview survey administered in 1973 and 1974 to residents of the Ottawa-Carleton, Ontario and

<sup>1</sup>The research was accomplished while this author was on leave at Research Laboratories, General Motors Corporation.

<sup>2</sup>The authors acknowledge the generous cooperation of the staff of the Regional Municipality of Ottawa-Carleton in supplying the data and providing valuable comments regarding the research reported herein.

Outaouais, Quebec Regions surrounding and including the Canadian National Capital. The total sample employed herein included 543 persons who reported making a regularly scheduled work trip. These persons reported their perceptions of the times and costs involved in making their last work trip by automobile and by bus and ranked these two modal choice alternatives on the twenty-five attributes listed in Table 1. The rankings were solic-

TABLE 1

ATTRIBUTE AS WORDED ON ATTITUDINAL QUESTIONNAIRE	ABBREVIATION
COMFORTABLE SEATING	SEATING COMFORT
DEPENDABILITY OF ON-TIME ARRIVAL	DEPENDABLE ARRIVAL TIME
AVAILABILITY MORE OR LESS WHEN YOU WANT IT	AVAILABILITY
ATTRACTIVENESS OF VEHICLE	VEHICLE ATTRACTIVENESS
LOW NOISE LEVEL IN VEHICLE	VEHICLE NOISE
VEHICLE SAFETY	VEHICLE SAFETY
SMOOTHNESS OF RIDE	RIDE SMOOTHNESS
PRIVACY FROM OTHER PEOPLE	PRIVACY
AVOIDING EXPOSURE TO TRAFFIC CONGESTION	TRAFFIC CONGESTION
MINIMUM EXPOSURE TO BODILY CROWDING	BODILY CROWDING
LOW OUT-OF-POCKET COST	OUT-OF-POCKET COST
LOW RIDING TIME	RIDING TIME
LOW WALKING TIME	WALKING TIME
LOW WAITING TIME	WAITING TIME
OPPORTUNITY TO MEET AND TALK TO OTHER PEOPLE	OPPORTUNITY TO MEET OTHERS
OPPORTUNITY TO RELAX	OPPORTUNITY TO RELAX
OPPORTUNITY TO READ	OPPORTUNITY TO READ
CONTINUOUS RIDE, FEW STOPS	RIDE CONTINUITY
PROTECTION FROM WEATHER ON ENTIRE TRIP	WEATHER PROTECTION
FLEXIBLE DESTINATION, CAN GO ANYWHERE	FLEXIBLE DESTINATION
NOT HAVING TO CHANGE VEHICLES	VEHICLE TRANSFER
YEAR-ROUND TEMPERATURE COMFORT IN VEHICLE	VEHICLE TEMPERATURE
ASSURANCE OF HAVING A SEAT	SEAT ASSURANCE
SECURITY FROM UNDESIRABLE ACTS OF OTHERS	PERSONAL SECURITY
MINIMUM POLLUTION PER PERSON CARRIED	POLLUTION PER PERSON

ited as perceived satisfactions based on a six-point semantic differential scale ranging from "very satisfied" to "very dissatisfied." The exact questions are detailed in Recker and Golob (1975).

Choice Constraint Segmentation

Three dimensions of perceptions by survey respondents of the accessibilities of the bus and auto to them were employed to capture supply-side constraints. Respondents' answers to questions regarding their perceived access to automobile were used to formulate a scale of automobile accessibility. The scale points, assumed to be at equal intervals, are: 1) no access even as passenger, 2) occasional access as passenger, 3) occasional access as driver, 4) access anytime as passenger, 5) access anytime as driver, and 6) ownership of personal automobile. Two ratio scales of bus accessibility were formulated by directly translating firstly the sum of each respondent's perceived waiting time and walking time and secondly

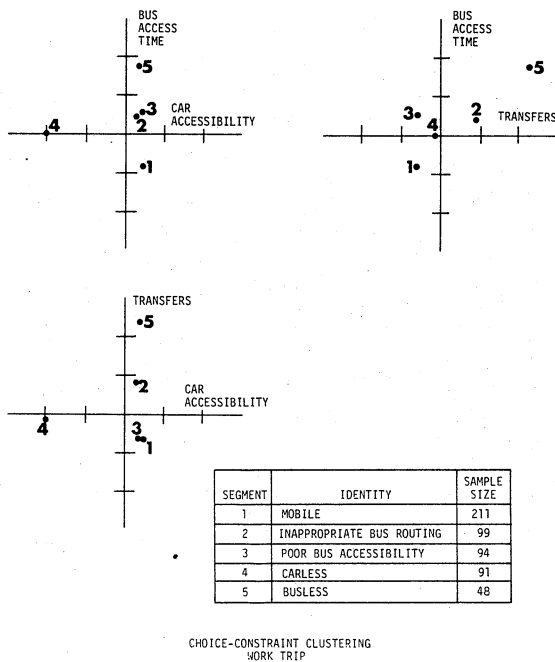
the number of transfers required to make the work trip by bus.

A clustering procedure was used to identify subgroups of the sample population which are relatively homogeneous with respect to these variables used to measure the accessibility of the alternatives. The variables are standardized to zero mean and unit variance to eliminate clustering bias due to scale differences. From the large number of clustering techniques available to accomplish such a segmentation, the technique chosen as representing a good compromise between sophistication and computational complexity was the ISODATA algorithm of Ball and Hall (1967). The number of clusters was determined by an iterative search process employing as a criterion a pseudo F-ratio of the total between-group variance divided by its degrees of freedom to the pooled within-group variance divided by its degrees of freedom.

The 543 respondents were found to be best segmented into five clusters. The stability of the resulting clusters was confirmed by performing sensitivity analyses with respect to cluster centers.

The positions of the five group centers in the three-dimensional standardized space were used to interpret the segments (Figure 1). Segment 1, labeled the "Mo-

FIGURE 1



bile" segment, is composed of individuals who are in the relatively most favorable position with respect to their modal choice alternatives. Segment 2, labeled the "Inappropriate Bus Routing" segment, includes individuals with high accessibility to the auto, bus access time slightly greater than the sample mean and number of transfers required for the work trip almost one standard deviation greater than the mean. Individuals in segment 3 the "Poor Bus Accessibility" segment, share the "Inappropriate Bus Routing" segment's problems of accessibility to the bus. However, unlike the "Inappropriate Bus Routing" segment, once the bus is accessed individuals within the "Poor Bus Accessibility" segment can reach their respective work locations with few, if any, transfers. Segment 4, the "Carless" segment, is differentiated from all other segments by their relatively

low accessibility to the auto. Segment 5, the "Busless" segment, is differentiated from all other groups by their relative lack of bus service.

#### Travelers' Perceptions of Choice Alternatives

Individuals' perceptions of their modal alternatives were analyzed by determining the factor structure of respondents' attitudes toward the set of twenty-five descriptive attributes of auto and bus shown in Table 1. Since it was hypothesized that individuals faced with different supply-side constraints might view their choice alternatives differently, separate factor analyses were performed for the total sample, for the "Mobile" segment and for a sample consisting of the remaining subgroups identified through the choice-constraint cluster analyses, labeled the complement to "Mobile" segment. Separate factor analyses of each of the subgroups within the complement to "Mobile" segment was not attempted due to sample size restrictions. Principal components analysis and varimax orthogonal rotation were employed.

Attributes included in the factor analyses were determined by an iterative process in which attributes having low correlations (factor loadings) with each of the factors retained, or those not having a single dominant correlation, were deleted from the correlation matrices. The adjusted matrices were then refactored until only attributes having significantly high loadings in a single factor remain. Attributes which were included in the final factors obtained using this procedure were termed "factorable" attributes. Attributes deleted by this process were termed "non-factorable" attributes. The number of factors retained for each analysis was determined by comparing eigenvalues with those obtained from analyses of random data matrices of the same order as the actual data matrices and the "Kaiser rule," in which all eigenvalues  $\geq 1$  are retained. (Where these two criteria did not result in selection of the same number of factors, selection between the two criteria was made by subjective judgment based on ease of interpretation provided by the factor loadings.)

A comparison of factor structures for the two modal choice alternatives was performed to determine the degree of similarity in perceptions. In cases in which the set of factorable variables was common to both auto and bus, the comparison was made by the orthogonal rotation procedure proposed by Gensch and Golob (1975). In cases in which the set of factorable variables was not common, this analytic procedure was replaced with a subjective similarity judgement based on inspection of factor loadings for the common attributes. Factors which were found to be similar across both alternatives were classified as "generic" factors. Factors which were unique to perception of a single choice alternative were classified as "alternative-specific" factors.

Factor analyses of the ratings of the twenty-five modal attributes listed in Table 1 for the auto and bus modes for the total sample of 543 respondents resulted in the selection of four factors to represent the bus attribute perceptions and five factors to represent the auto attribute perceptions. The set of "non-factorable" attributes is shown in Table 2. Table 3 shows the bus perception factors for the total sample. The "factor description" column gives a subjective label for each factor together with the percentage of the variance in the original attribute set which is accounted for by this factor. The "attribute" column lists the attributes with loadings on each factor which are significantly different from zero, and the "percent variance" columns list the percent variance of each attribute which is accounted for by the factor in question and the percent which is accounted for by all other factors combined.



TABLE 2

TOTAL SAMPLE: WORK TRIP NON-FACTORABLE ATTRIBUTES	
BUS (8)	AUTO (8)
SEATING COMFORT TRAFFIC CONGESTION OUT-OF-POCKET COST WEATHER PROTECTION PERSONAL SECURITY POLLUTION PER PERSON BODILY CROWDING VEHICLE TEMPERATURE	SEATING COMFORT TRAFFIC CONGESTION OUT-OF-POCKET COST WEATHER PROTECTION PERSONAL SECURITY POLLUTION PER PERSON BODILY CROWDING VEHICLE TEMPERATURE

TABLE 3

TOTAL SAMPLE : BUS WORK TRIP FACTORS

FACTOR DESCRIPTION (% VARIANCE EXPLAINED)	ATTRIBUTES INCLUDED IN FACTOR			
	ATTRIBUTE	FACTOR LOADING	PERCENT VARIANCE	
			IN FACTOR	IN ALL OTHER FACTORS
SERVICE (27%)	WAITING TIME	.82	67	04
	RIDING TIME	.75	56	05
	VEHICLE TRANSFER	.74	55	06
	AVAILABILITY	.69	48	15
	DEPENDABLE ARRIVAL TIME	.69	47	16
	RIDE CONTINUITY	.69	47	14
	WALKING TIME	.68	46	08
VEHICLE RIDE QUALITY (15%)	FLEXIBLE DESTINATION	.67	44	10
	RIDE SMOOTHNESS	.77	59	11
	VEHICLE SAFETY	.75	56	04
	VEHICLE NOISE	.72	52	14
PERSONAL AUTONOMY (12%)	VEHICLE ATTRACTIVENESS	.63	40	29
	OPPORTUNITY TO READ	.80	63	08
	OPPORTUNITY TO RELAX	.78	62	14
SOCIALITY (08%)	SEAT ASSURANCE	.57	32	23
	OPPORTUNITY TO MEET OTHERS	.72	51	09
	VEHICLE ATTRACTIVENESS	.53	29	40
	PRIVACY	.52	27	21

These two pieces of information (which sum to the communality) indicate respectively the strength and uniqueness of the attribute-factor relationship. The auto perception factors are presented in Table 4. Here the

TABLE 4

TOTAL SAMPLE : AUTO WORK TRIP FACTORS

FACTOR DESCRIPTION (% VARIANCE EXPLAINED)	ATTRIBUTES INCLUDED IN FACTOR			
	ATTRIBUTE	FACTOR LOADING	PERCENT VARIANCE	
			IN FACTOR	IN ALL OTHER FACTORS
CONVENIENCE (15%)	VEHICLE TRANSFER	.79	62	03
	SEAT ASSURANCE	.72	52	04
	WAITING TIME	.64	41	16
	FLEXIBLE DESTINATION	.62	38	24
	WALKING TIME	.59	35	20
VEHICLE QUALITY (13%)	VEHICLE NOISE	.82	67	04
	RIDE SMOOTHNESS	.77	59	09
	VEHICLE ATTRACTIVENESS	.68	46	15
	PRIVACY	.56	31	18
PERFORMANCE (12%)	RIDING TIME	.79	62	06
	RIDE CONTINUITY	.66	49	12
	DEPENDABLE ARRIVAL TIME	.59	35	18
PERSONAL ENVIRONMENT (10%)	OPPORTUNITY TO MEET OTHERS	.75	56	03
	OPPORTUNITY TO READ	.74	55	04
	OPPORTUNITY TO RELAX	.66	44	16
AVAILABILITY (08%)	AVAILABILITY	.72	52	11
	VEHICLE SAFETY	.60	36	31

most dominant factor, "convenience," does not explain as great a percentage of attribute variance as does the dominant bus "service" factor.

None of the bus and auto factors were determined to be directly compatible, and consequently there were no

generic factors. For the total sample, the attributes that make up the bus "service" factor split into the two auto factors "convenience" and "performance." Moreover, subtle differences exist both between the auto "vehicle quality" factor which includes the "privacy" attribute and the bus "vehicle ride quality" factor which includes the "vehicle safety" attribute, as well as between the auto "personal environment" factor which includes the "opportunity to meet others" attribute and the bus "personal autonomy" factor which includes the "seat assurance" attribute. In addition, perception of auto includes a dimension linking its availability to safety while perception of bus includes a dimension associated with social interactions.

Factor analyses of the attribute satisfaction ratings for the bus and auto modes by the 211 respondents classified into the "mobile" market segment resulted in the selection of four factors to represent perceptions of the bus mode and five factors to represent perceptions of the auto mode. Differences between the total sample and the "mobile" segment perceptions are apparent when comparing Tables 5, 6 and 7 with Tables 2, 3 and 4.

TABLE 5

"MOBILE" SEGMENT : WORK TRIP NON-FACTORABLE ATTRIBUTES	
BUS (8)	AUTO (11)
SEATING COMFORT	SEATING COMFORT
VEHICLE ATTRACTIVENESS	VEHICLE ATTRACTIVENESS
PERSONAL SECURITY	PERSONAL SECURITY
POLLUTION PER PERSON	POLLUTION PER PERSON
OUT-OF-POCKET COST	DEPENDABLE ARRIVAL TIME
WEATHER PROTECTION	AVAILABILITY
SEAT ASSURANCE	PRIVACY
OPPORTUNITY TO RELAX	WAITING TIME
	FLEXIBLE DESTINATION
	VEHICLE TEMPERATURE
	VEHICLE SAFETY

TABLE 6

"MOBILE" SEGMENT : BUS WORK TRIP FACTORS

FACTOR DESCRIPTION (% VARIANCE EXPLAINED)	ATTRIBUTES INCLUDED IN FACTOR			
	ATTRIBUTE	FACTOR LOADING	PERCENT VARIANCE	
			IN FACTOR	IN ALL OTHER FACTORS
SERVICE (26%)	WAITING TIME	.84	70	07
	VEHICLE TRANSFER	.80	64	02
	AVAILABILITY	.76	59	04
	RIDE CONTINUITY	.74	54	14
	RIDING TIME	.71	50	15
	FLEXIBLE DESTINATION	.71	50	07
	WALKING TIME	.70	50	00
VEHICLE RIDE QUALITY (14%)	DEPENDABLE ARRIVAL TIME	.67	59	04
	VEHICLE SAFETY	.82	67	00
	RIDE SMOOTHNESS	.75	57	08
CROWDING (13%)	VEHICLE NOISE	.60	36	15
	BODILY CROWDING	.81	65	08
	PRIVACY	.69	48	26
PERSONAL ENVIRONMENT (09%)	TRAFFIC CONGESTION	.69	48	12
	OPPORTUNITY TO MEET OTHERS	.81	66	01
	OPPORTUNITY TO READ	.73	54	05
	VEHICLE TEMPERATURE	.66	43	14

Individuals in the "mobile" sample, expectedly, are more sensitive to differences in meaning among many of the

TABLE 7

"MOBILE" SEGMENT : AUTO WORK TRIP FACTORS				
FACTOR DESCRIPTION (% VARIANCE EXPLAINED)	ATTRIBUTE	FACTOR LOADING	PERCENT VARIANCE	
			IN FACTOR	IN ALL OTHER FACTORS
			CONVENIENCE (17%)	VEHICLE TRANSFER WEATHER PROTECTION WALKING TIME SEAT ASSURANCE
PERFORMANCE (14%)	RIDING TIME TRAFFIC CONGESTION OUT-OF-POCKET COST RIDE CONTINUITY	.72 .71 .63 .62	52 50 40 38	12 17 20 19
PERSONAL ENVIRONMENT (12%)	OPPORTUNITY TO READ OPPORTUNITY TO MEET OTHERS OPPORTUNITY TO RELAX	.79 .72 .70	62 52 49	03 03 13
VEHICLE RIDE QUALITY (11%)	VEHICLE NOISE RIDE SMOOTHNESS	.85 .77	72 59	07 13
CROWDING (10%)	BODILY CROWDING	.78	62	06

attributes of auto which are confounded by the total sample and thus there are more non-factorable attributes for auto in Table 5 than in Table 2. Comparing the bus perception factors for the two groups (Tables 3 and 6), the "service" and "vehicle ride quality" factors for the "mobile" segment and the total sample are quite similar, but significant differences in perception exist with respect to the other two factors. For the "mobile" segment these two dimensions involve aspects of vehicular and bodily crowding and the personal environment of the passenger, respectively. For the total sample these two dimensions are associated with preserving personal autonomy and with social interactions, respectively. Comparing Tables 4 and 7, the "personal environment" factors for the two segments are similar as are the "performance" and "convenience" factors. The "vehicle quality" factor for the total sample is more concisely defined by the "mobile" segment as a "vehicle ride quality" factor, and the factor pairing "availability" and "vehicle safety" attributes for the total sample has been replaced by the more precise "crowding" factor for the "mobile" segment.

Factor analyses for the complement to "mobile" segment (which includes the "carless," "busless," "poor bus accessibility" and "inappropriate bus routing" segments) resulted in the use of one less factor to describe the latent structure of both the bus and auto perceptions. Moreover, the set of "non-factorable" attributes obtained for the complement to "mobile" segment (Table 8)

TABLE 8

COMPLEMENT TO "MOBILE" SEGMENT : WORK TRIP NON-FACTORABLE ATTRIBUTES	
BUS (7)	AUTO (7)
BODILY CROWDING OUT-OF-POCKET COST VEHICLE TEMPERATURE POLLUTION PER PERSON	BODILY CROWDING OUT-OF-POCKET COST VEHICLE TEMPERATURE POLLUTION PER PERSON
-----	-----
SEATING COMFORT TRAFFIC CONGESTION OPPORTUNITY TO MEET OTHERS	AVAILABILITY VEHICLE SAFETY PERSONAL SECURITY

bears little resemblance to that obtained for the "mobile" segment (Table 5). In the case of the bus perception factors, comparison of Tables 9 and 6 reveals that the service factor accounting for the greatest propor-

TABLE 9

COMPLEMENT TO "MOBILE" SEGMENT : BUS WORK TRIP FACTORS

FACTOR DESCRIPTION (% VARIANCE EXPLAINED)	ATTRIBUTE	FACTOR LOADING	PERCENT VARIANCE	
			IN FACTOR	IN ALL OTHER FACTORS
			SERVICE (25%)	WAITING TIME RIDING TIME AVAILABILITY DEPENDABLE ARRIVAL TIME FLEXIBLE DESTINATION VEHICLE TRANSFER RIDE CONTINUITY WALKING TIME WEATHER PROTECTION
VEHICLE QUALITY (16%)	VEHICLE ATTRACTIVENESS RIDE SMOOTHNESS VEHICLE NOISE VEHICLE SAFETY	.76 .75 .74 .73	58 56 55 53	01 14 15 05
PERSONAL ENVIRONMENT AND AUTONOMY (15%)	OPPORTUNITY TO READ OPPORTUNITY TO RELAX SEAT ASSURANCE PRIVACY PERSONAL SECURITY	.78 .75 .58 .56 .53	61 56 34 31 28	00 11 18 14 20

tion of variance in the attribute ratings is essentially the same for the "mobile" and complement to "mobile" segments. The vehicle ride quality factor of the "mobile" segment is closely related to the vehicle quality factor for the complement to "mobile" segment, but the latter factor includes "vehicle attractiveness," a unique or unfavorable attribute for the "mobile" segment. Finally, while two factors, crowding and sociability, account for the remaining linear interdependencies for the "mobile" segment bus perceptions, only a single factor, personal environment and autonomy, accounts for the perceptions not described by the service and vehicle quality factors for the complement. The inability to separate perceptions of the five attributes comprising the personal environment and autonomy factor may be a result of confounding different patterns of perception in a single composite sample. In the case of the auto perception factors (Tables 10 and 7), the personal environ-

TABLE 10

COMPLEMENT TO "MOBILE" SEGMENT : AUTO WORK TRIP FACTORS

FACTOR DESCRIPTION (% VARIANCE EXPLAINED)	ATTRIBUTE	FACTOR LOADING	PERCENT VARIANCE	
			IN FACTOR	IN ALL OTHER FACTORS
			CONVENIENCE (19%)	VEHICLE TRANSFER WEATHER PROTECTION SEAT ASSURANCE WAITING TIME FLEXIBLE DESTINATION WALKING TIME
VEHICLE QUALITY (16%)	RIDE SMOOTHNESS VEHICLE NOISE VEHICLE ATTRACTIVENESS SEATING COMFORT	.81 .80 .70 .65	65 63 49 43	05 02 13 11
PERFORMANCE (14%)	RIDING TIME DEPENDABLE ARRIVAL TIME TRAFFIC CONGESTION RIDE CONTINUITY	.77 .72 .71 .63	59 52 50 40	05 06 04 22
PERSONAL ENVIRONMENT (10%)	OPPORTUNITY TO READ OPPORTUNITY TO MEET OTHERS OPPORTUNITY TO RELAX	.79 .77 .62	62 59 39	02 01 21

ment, convenience, and performance latent perception factors for the complement to "mobile" segment are approximately the same, but the "mobile" segment vehicle ride quality factor is expanded to include "vehicle attractiveness" and "seating comfort" for the complement to "mobile" segment.

The results of these factor analyses indicate that supply-side constraints are significantly related to perception of modal choice alternatives. In general, it can be stated that segmentation according to supply-side constraint conditions leads to more sharply defined perceptual attribute spaces than those obtained using the non-segmented sample.

### Choice Models

To investigate relationships between choice constraints and decision-making behavior, attitudinal modal choice models were developed and estimated for the total sample and for four of the five segments identified by the cluster analyses. (The sample size of the "busless" segment was judged to be too small to permit choice model parameter estimation.)

It is hypothesized that an individual decision maker's overall preference ranking of a choice alternative is a function of the utility which that alternative holds for the individual, specified in terms of individual  $i$ 's attitudes toward alternative  $k$ . A utility form which is linear and additive in terms of attitudes toward the attributes of the alternatives is assumed (Wilkie and Pessmier, 1973).

In light of the division of perception toward each alternative  $k$  into a set of non-factorable attributes,  $S_N^k$ , and a set of latent factors,  $Q^k$ , utility is here specified by

$$U_i^k = \sum_{j \in S_N^k} a_j^k x_{ij}^k + \sum_{q \in Q^k} a_q^k y_{iq}^k + \xi_i^k \quad (1)$$

where

- $U_i^k$  = utility of alternative  $k$  to individual  $i$ ,
- $x_{ij}^k$  = manifest rating by individual  $i$  of alternative  $k$  on attribute  $j$ ,
- $y_{iq}^k$  = latent (i.e., unobserved) scores for alternative  $k$  on factor  $q$  for individual  $i$ ,
- $a_j^k$  = utility weight reflecting the importance of the  $j$ th attribute in contributing to the overall utility of alternative  $k$  to individual  $i$ ,
- $a_q^k$  = utility weight reflecting the similar importance of the  $q$ th latent factor, and
- $\xi_i^k$  = random component, assumed to be independent and identically distributed across all individuals.

The utility weights  $a_j^k$  and  $a_q^k$  are assumed to be invariant across individuals in a particular market segment.

Planners using the models for forecasting purposes usually think in terms of attributes and not in terms of linear composites of attributes even if these composites represent psychological dimensions of perception. Thus, to obtain a model structure more useful for interpretation and prediction, the latent factors in Equation (1), can be approximated by representative attributes with high loadings on these factors. Thus, equation (1) is rewritten as:

$$U_i^k = \sum_{j \in S_N^k} a_j^k x_{ij}^k + \sum_{j \in S_F^{k*}} b_j^k x_{ij}^k + \epsilon_i^k \quad (2)$$

or

$$U_i^k = v_i^k + \epsilon_i^k \quad (3)$$

where

- $b_j^k$  = a modified utility weight (an equation which is given by Recker and Golob, 1975),

$S_F^{k*}$  = the set of attributes chosen each to represent one and only one of the factors in set  $Q^k$  describing perception of alternative  $k$ ,

$\epsilon_i^k$  = an error term representing the  $\xi_i^k$  error term of equation (1) and errors introduced by approximating latent factors by attributes, and

$v_i^k$  = deterministic utility component.

Operationally, the set  $S_F^{k*}$  can be specified as being comprised of the attributes  $j^*$  such that  $f_{j^*q}$  is maximum over all  $j \in S_F^k$ , for each  $q \in Q^k$ . However, this choice is somewhat arbitrary if for some  $q$  there are two or more attributes with factor loadings which are approximately equal and of high absolute value. The choice criteria for establishing  $S_F^{k*}$  can then be related to specific planning objectives. This is judged to be a definite advantage of the present methodology: it can potentially be used to test a variety of policy issues without change in analytical structure. It is the antithesis of "single model" methodologies represented in an extreme case by the use of step-wise or screening linear regressions to find "optimal" subsets of independent variables.

The probability that individual  $i$  will prefer alternative  $k$  from a set of available alternatives  $A$ , denoted by  $P_i(k:A)$ , can be written in terms of the simplified utility of expression (3) as

$$P_i(k:A) = \text{Prob} \{ (v_i^k + \epsilon_i^k) > (v_i^l + \epsilon_i^l) \text{ for all } l \in A, l \neq k \}. \quad (4)$$

Or,

$$P_i(k:A) = \text{Prob} \{ (v_i^k - v_i^l) > (\epsilon_i^l - \epsilon_i^k) \text{ for all } l \in A, l \neq k \}. \quad (5)$$

If it is assumed that consumers adjust their travel-related behavior so as to maximize their utilities, Gumbel (1954) has shown that for a wide range of distributions of random utility variables, the random terms  $\epsilon_i^k$  are asymptotically independently identically distributed with the Weibull (Gnedenko extreme value) distribution. The choice probability then takes the form

$$P_i(k:A) = 1 / (1 + \sum_{\substack{l \in A \\ l \neq k}} \exp (v_i^k - v_i^l)). \quad (6)$$

A detailed derivation of this type of strict-utility model, called the multinomial logit model, is provided by McFadden (1973).

The form of the choice model employed herein is thus specified by substituting the deterministic ( $v_i^k$ ) component of  $U_i^k$  given by equation (2) into equation (6), for  $k \in A$ . The dependent variable is the observed choice, where  $P_i(k:A)$  takes the value 1 when  $k$  is chosen, 0 when  $k$  is not chosen. Parameters to be estimated are the utility weights  $a_j^k$  for all attributes in set  $S_N^k$  and  $b_j^k$  for all attributes in set  $S_F^{k*}$ . For "generic" attributes these weights are assumed to be mode-independent ( $a_j$  and  $b_j$ ). The parameters were estimated using maximum likelihood techniques (McFadden, 1968).

It is thus hypothesized that individuals compare pairs of choice alternatives on the basis of absolute levels of their perceived satisfactions with alternative-specific attributes compensatory with perceived differences in satisfactions with generic or choice-independent attributes. This says that decision makers will make relative comparisons of alternatives on all attributes which have consistent meaning for each alternative, but will make absolute evaluations of the alternatives on

all attributes which have unique meanings. Such a conceptualization considers attitudinal preference model specifications in which all evaluations between alternatives are treated as differences between attribute scores (e.g., Hansen, 1969) as a special case.

Because of the lack of goodness-of-fit measures with well-defined statistical properties (such as the linear regression coefficient of determination,  $R^2$ ) for such probabilistic choice models, emphasis was placed on coefficient significance tests and on predictive performance criteria for evaluating the models. One such indicator is the ratio of choices predicted correctly by the models; this is determined as the ratio of the number of times the predicted probability of the chosen alternative is greater than that of a non-chosen alternative. This ratio was also disaggregated by alternative chosen (bus or auto). In addition, two different measures (both termed "pseudo  $R^2$ ", or  $\rho^2$ ) which are non-linear analogies of the linear  $R^2$  measure were also used to evaluate overall model performance. These measures were proposed by McFadden (1968) and Cragg (1968) (this latter measure is attributed to Theil) and are detailed in Burns, Golob and Nicolaidis (1975). Unfortunately, distributional properties have not been defined for either  $\rho^2$  measure and maximum values of the  $\rho^2$  measures can be less than 1.0.

To determine the sensitivities of the choice probabilities to changes in the various attributes affecting modal choice elasticities were calculated and compared. Aggregate elasticities are developed from the individual elasticities associated with equation (6) to estimate the overall sensitivities of choice probabilities to uniform percent changes in explanatory variables for all individuals (Recker and Golob, 1975).

Results of the maximum likelihood estimation of a logit choice model of equation (6) for the total sample are displayed in Table 11. Presented are the coefficients

TABLE 11

TOTAL SAMPLE : WORK TRIP  
LOGIT CHOICE MODEL  
(SAMPLE SIZE = 543)

VARIABLE	FACTOR	COEFFICIENT	t
BUS-WALKING TIME	BUS-SERVICE	0.657	7.08
AUTO-AVAILABILITY	AUTO-AVAILABILITY	0.478	4.47
OUT-OF-POCKET COST		0.206	3.51
AUTO-WALKING TIME	AUTO-CONVENIENCE	0.289	2.95
TRAFFIC CONGESTION		0.114	1.84

McFadden  $\rho^2 = 0.394$  Theil  $\rho^2 = 0.421$

PERCENT OF CHOICES PREDICTED CORRECTLY	= 31.2%
PERCENT OF CHOICES PREDICTED CORRECTLY WHEN AUTO WAS THE CHOSEN MODE (78.2% OF TOTAL CASES)	= 95.4%
PERCENT OF CHOICES PREDICTED CORRECTLY WHEN BUS WAS THE CHOSEN MODE (21.8% OF TOTAL CASES)	= 30.2%

for all attributes for which the coefficients are significantly different from zero at the 95% confidence level. All attributes with coefficients insignificantly different from zero at this confidence level were not included in the estimations. The coefficient values are listed in order of their t-statistics which are asymptotically distributed as t-statistics in a linear model (Theil, 1971).

Of those factors in the logit choice model estimation for the total sample included, the bus-service factor (represented by the "walking time" attribute) exhibited the greatest explanatory power as indicated by coefficient t-statistics (Table 11). This factor was followed approximately in explanatory power by the auto availability factor, the generic attribute "out-of-pocket cost," the auto-convenience factor, and the generic attribute "traffic congestion". The interpretation is that the level of service provided by the bus system, as measured by attributes such as "vehicle transfers," "seat assurance," "waiting time," "flexible destination," and "walking time," explains choice between auto and bus for the work trip.

Aggregate elasticity estimates are shown in Table 12 for

TABLE 12

TOTAL SAMPLE : WORK TRIP  
AGGREGATE LOGIT CHOICE ELASTICITIES  
(SAMPLE SIZE : 543)

VARIABLE	FACTOR	AUTO ALTERNATIVE		BUS ALTERNATIVE	
		ELASTICITY	CROSS-ELASTICITY	ELASTICITY	CROSS-ELASTICITY
BUS-WALKING TIME	BUS-SERVICE		-1.366	1.864	
AUTO-AVAILABILITY	AUTO-AVAILABILITY	0.449			-2.247
OUT-OF-POCKET COST		0.146	-0.420	0.595	-0.773
AUTO-WALKING TIME	AUTO-CONVENIENCE	0.259			-1.328
TRAFFIC CONGESTION		0.078	-0.172	0.250	-0.401

the total work trip sample. The probability of choosing the bus alternative was found to be most sensitive to respondents' evaluations of the availability of the auto alternative; a ten percent decline in satisfactions with auto availability would be expected to result in over a twenty percent increase in the probability of choosing bus. Other explanatory variables exhibiting relatively high elasticities for choice of bus include the bus "walking time" attribute representing the bus service factor and the auto "walking time" attribute representing the auto convenience factor. The only similarly high elasticity on choice of auto is that associated with the bus "walking time" attribute.

To demonstrate the latitude available to the analyst in the selection of the attributes to represent the latent perception factors which are found to be significant explanatory variables, Tables 13 and 14 list results from estimation of a choice model using "vehicle transfer" as a substitute for "walking time," to represent the bus service factor. Results from these two versions of the choice model (and other versions tested but not reported herein) are judged to be approximately the same. Motivations for choosing any particular representative attribute would be related to objectives of testing policy alternatives in planning applications of the model.

Results from logit calibrations for the "mobile" segment are summarized in Tables 15 and 16. The goodness-of-fit measures in Table 15 indicate a slightly better explanation of choice than that obtained for the total work trip sample. These results are similar to those for the total sample in that both indicate bus-service and auto-convenience factors as well as auto-traffic congestion (as representative of the auto-performance factor for the "mobile" segment) to be determinant to choice of mode. However, whereas auto-availability and out-of-pocket cost are determined as important to choice for the total sample, analysis of the more homogeneous "mobile" segment of the total population indicates that

TABLE 13

TOTAL SAMPLE : WORK TRIP  
VERSION 2 - LOGIT CHOICE MODEL

VARIABLE	FACTOR	COEFFICIENT	t
BUS-VEHICLE TRANSFER	BUS-SERVICE	0.453	6.51
AUTO-AVAILABILITY	AUTO-AVAILABILITY	0.335	3.53
OUT-OF-POCKET COST		0.244	4.13
AUTO-WALKING TIME	AUTO-CONVENIENCE	0.242	2.61
TRAFFIC CONGESTION		0.106	1.74

McFadden  $\rho^2 = 0.367$  Theil  $\rho^2 = 0.399$ 

PERCENT OF CHOICES PREDICTED CORRECTLY	= 80.6%
PERCENT OF CHOICES PREDICTED CORRECTLY WHEN AUTO WAS THE CHOSEN MODE (78.2% OF TOTAL CASES)	= 96.6%
PERCENT OF CHOICES PREDICTED CORRECTLY WHEN BUS WAS THE CHOSEN MODE (21.8% OF TOTAL CASES)	= 23.9%

TABLE 14

TOTAL SAMPLE : WORK TRIP  
VERSION 2  
AGGREGATE LOGIT CHOICE ELASTICITIES

VARIABLE	FACTOR	AUTO ALTERNATIVE		BUS ALTERNATIVE	
		ELASTICITY	CROSS-ELASTICITY	ELASTICITY	CROSS-ELASTICITY
BUS-VEHICLE TRANSFER	BUS-SERVICE		-0.809	1.237	
AUTO-AVAILABILITY	AUTO-AVAILABILITY	0.327			-1.562
OUT-OF-POCKET COST		0.180	-0.460	0.732	-0.905
AUTO-WALKING TIME	AUTO-CONVENIENCE	0.227			-1.103
TRAFFIC CONGESTION		0.076	-0.147	0.243	-0.368

TABLE 15

"MOBILE" SEGMENT : WORK TRIP  
LOGIT CHOICE MODEL  
(SAMPLE SIZE = 211)

VARIABLE	FACTOR	COEFFICIENT	t
BUS-AVAILABILITY	BUS-SERVICE	0.890	5.41
AUTO-WALKING TIME	AUTO-CONVENIENCE	0.473	3.49
AUTO-TRAFFIC CONGESTION	AUTO-PERFORMANCE	0.387	2.69
AUTO-OPPORTUNITY TO READ	AUTO-PERSONAL ENVIRONMENT	0.267	2.08

McFadden  $\rho^2 = 0.471$  Theil  $\rho^2 = 0.479$ 

PERCENT OF CHOICES PREDICTED CORRECTLY	= 82.8%
PERCENT OF CHOICES PREDICTED CORRECTLY WHEN AUTO WAS THE CHOSEN MODE (79.2% OF TOTAL CASES)	= 94.6%
PERCENT OF CHOICES PREDICTED CORRECTLY WHEN BUS WAS THE CHOSEN MODE (20.8% OF TOTAL CASES)	= 38.5%

for this most unconstrained group these aspects are not significant descriptors of modal choice. As shown in Table 16, the auto convenience factor represented by the "walking time" attribute was most strongly related to the "mobile" segments' choice of the bus mode. This cross-elasticity of -2.225 can be compared to the cor-

TABLE 16

"MOBILE" SEGMENT : WORK TRIP  
AGGREGATE LOGIT CHOICE ELASTICITIES  
(SAMPLE SIZE = 211)

VARIABLE	FACTOR	AUTO ALTERNATIVE		BUS ALTERNATIVE	
		ELASTICITY	CROSS-ELASTICITY	ELASTICITY	CROSS-ELASTICITY
BUS-AVAILABILITY	BUS-SERVICE		-2.040	2.022	
AUTO-WALKING TIME	AUTO-CONVENIENCE	0.355			-2.225
AUTO-TRAFFIC CONGESTION	AUTO-PERFORMANCE	0.227			-1.455
AUTO-OPPORTUNITY TO READ	AUTO-PERSONAL ENVIRONMENT	0.102			-0.690

responding value of -1.328 obtained for the total sample. The revealed sensitivity of modal choice to auto performance is an intuitively satisfying result; individuals in the "mobile" segment would be expected to be critical in their evaluation of auto, since bus is accessible and is thus a viable alternative to their travel by auto.

The results of the logit estimation for the "inappropriate bus routing" segment are shown in Table 17. The

TABLE 17

"INAPPROPRIATE BUS ROUTING" SEGMENT : WORK TRIP  
LOGIT CHOICE MODEL  
(SAMPLE SIZE = 99)

VARIABLE	FACTOR	COEFFICIENT	t
AUTO-VEHICLE TRANSFER	AUTO-CONVENIENCE	1.217	3.88
BUS-DEPENDABLE ARRIVAL TIME	BUS-SERVICE	1.072	3.09
OUT-OF-POCKET COST		0.444	1.99

McFadden  $\rho^2 = 0.660$  Theil  $\rho^2 = 0.600$ 

PERCENT OF CHOICES PREDICTED CORRECTLY	= 90.8%
PERCENT OF CHOICES PREDICTED CORRECTLY WHEN AUTO WAS THE CHOSEN MODE (87.8% OF TOTAL CASES)	= 98.8%
PERCENT OF CHOICES PREDICTED CORRECTLY WHEN BUS WAS THE CHOSEN MODE (12.2% OF TOTAL CASES)	= 33.3%

goodness-of-fit measures indicate explanatory power significantly greater than the total sample and "mobile" segment models. The model is also slightly more parsimonious than the previous models, with only the auto convenience and bus service factors and the "out-of-pocket cost" attribute displaying coefficients significantly different from zero at the 95% confidence level. As with the total sample results, Table 17 indicates auto-convenience and bus-service factors and out-of-pocket cost are significant determinants of modal choice for the "inappropriate bus routing" segment. However, the auto-availability factor and "traffic congestion" attribute which were identified as significant for the total sample are not represented. Moreover, elasticity estimates (not shown for reasons of brevity) indicate significantly different sensitivities of choice between those determined for the total sample and "inappropriate bus routing" segment. For example, the sensitivity of choice of the bus mode to the variable representing the auto-convenience factor for the "inappropriate bus routing" segment was much greater.

The results of the logit estimation for the "poor bus accessibility" segment are displayed in Table 18. As in

TABLE 18

"POOR BUS ACCESSIBILITY" SEGMENT : WORK TRIP  
LOGIT CHOICE MODEL  
SAMPLE SIZE N = 94

VARIABLE	FACTOR	COEFFICIENT	t
OUT-OF-POCKET COST		0.912	3.03
BUS-WAITING TIME	BUS-SERVICE	0.578	2.09
CONSTANT		5.569	3.82

McFadden  $\rho^2 = 0.735$       Theil  $\rho^2 = 0.639$

PERCENT OF CHOICES PREDICTED CORRECTLY	= 92.5 %
PERCENT OF CHOICES PREDICTED CORRECTLY WHEN AUTO WAS THE CHOSEN MODE (91.4% OF TOTAL CASES)	= 97.6 %
PERCENT OF CHOICES PREDICTED CORRECTLY WHEN BUS WAS THE CHOSEN MODE (8.6% OF TOTAL CASES)	= 37.5 %

the case of the "inappropriate bus routing" model, the goodness-of-fit indices are significantly better than those obtained with the total sample. Also consistent with the previous model, the bus service factor and "out-of-pocket cost" were found to have significant explanatory power. This consistency between two segments faced with different aspects of bus supply-side problems is intuitively satisfying. Moreover, the inclusion of the auto convenience factor as a significant explanatory variable in the "inappropriate bus routing" model and the exclusion of this factor in the "poor bus accessibility" model emphasizes the differences in the bus supply-side problems faced by these two market segments. Individuals in the "inappropriate bus routing" segment can get to a bus but the service provided, as related to auto performance, might be unacceptable to them. Individuals in the "poor bus accessibility" segment, on the other hand, experience trouble in actually getting to and from the bus and thus have little reason to compare the bus with the performance of the automobile. Omitted factors related to problems of getting to and from the bus might account for the significance of the constant in the "poor bus accessibility" model. Comparing the aggregate elasticities for this segment (not shown) with those for the total sample, there is a much greater sensitivity to the generic attribute "out-of-pocket cost". This is a vivid example of information lost without segmentation.

The results for the "carless" segment are shown in Table 19. The goodness-of-fit is poorer than for any of the other modal choice models, although it is still judged to be very good for this class of models. Three of the four factors exhibiting significant explanatory power in this model are related to social and psychological factors involving the relationships between an individual and those inanimate objects and other persons immediately around him or her. Such diagnostic information is not obtainable from the choice model estimated for the total sample and provides clear evidence of the importance of choice-constraint segmentation. Moreover, the elasticities associated with the two attributes common to both the "carless" segment and total sample models differ greatly between models.

TABLE 19  
"CARLESS" SEGMENT : WORK TRIP  
LOGIT CHOICE MODEL  
(SAMPLE SIZE = 91)

VARIABLE	FACTOR	COEFFICIENT	t
BUS-WALKING TIME	BUS-SERVICE	0.777	3.34
AUTO-OPPORTUNITY TO READ	AUTO-PERSONAL ENVIRONMENT	0.486	2.80
BUS-OPPORTUNITY TO MEET OTHERS	BUS-SOCIABILITY	-0.682	-2.79
BUS-OPPORTUNITY TO READ	BUS-PERSONAL AUTONOMY	0.580	2.77
OUT-OF-POCKET COST		0.377	2.33

McFadden  $\rho^2 = 0.342$       Theil  $\rho^2 = 0.374$

PERCENT OF CHOICES PREDICTED CORRECTLY	= 75.9%
PERCENT OF CHOICES PREDICTED CORRECTLY WHEN AUTO WAS THE CHOSEN MODE (36.1% OF TOTAL CASES)	= 57.6%
PERCENT OF CHOICES PREDICTED CORRECTLY WHEN BUS WAS THE CHOSEN MODE (63.9% OF TOTAL CASES)	= 86.8%

The sample size for the "busless" segment was too small to permit model estimation.

#### Constraints as Conditional Probabilities

Consider the possibility that important causal relationships regarding choice of travel mode are indeed captured by the choice constraint segmentation, but differences between segments in terms of perception structure and attribute utility weights do not reflect all of the supply-side effects on actual choice behavior. Such a contention is consistent with general consumer behavior theories which identify situational variables intervening between preference and actual purchase or use (Sheth, 1975).

Assume that the probability that an individual from a particular market segment will choose alternative k is the product of the conditional probability that k is the preferred alternative given the accessibility of the alternative and the probability that k is accessible. Denote this accessibility probability by  $\gamma^k$  for each segment. The probability of an individual i in that segment choosing alternative k is then equal to the product  $\gamma^k P_i(k:A)$ , where  $P_i(k:A)$  is given by equation (6).

The modal accessibility probabilities  $\gamma^k$  were estimated for each of the four segments by determining a range of values of the ratio of  $\gamma^{BUS}$  to  $\gamma^{AUTO}$  which maximized the correct overall prediction ratio, subject to the condition that the prediction ratio for individuals choosing each of the two modes does not decrease. Results from the  $\gamma^k$  estimations are provided in Table 20. Incorporation of the accessibility probabilities led to improvements in only one of the four choice models. However, it is intuitively satisfying that this model was for the "poor bus accessibility" segment, which was the most difficult to describe using the logit preference model without accessibility modifications. The ranges of accessibility probability ratios for which the prediction ratios remain constant were also consistent with descriptive characteristics of the market segments. Since supply-side problems are minimal for the "mobile" segment, it is logical that it should have one of the tightest ranges. It is also consistent that the "carless" should be the only segment to have all accessibility probability ratios (bus to auto) greater than

TABLE 20

EFFECT OF  $\gamma^k$  ON PREDICTION RATIOS  
WORK TRIP

SEGMENT	CHOSEN MODE	$\frac{\gamma_{BUS}}{\gamma_{AUTO}}$	PREDICTION RATIO WITH $\gamma^k$	PREDICTION RATIO BY MODE WITH $\gamma^k$	PREDICTION RATIO WITHOUT $\gamma^k$	PREDICTION RATIO WITHOUT $\gamma^k$
MOBILE	AUTO BUS	.95-1.00	.828	.946 .385	.828	.946 .385
POOR BUS ACCESSIBILITY	AUTO BUS	.68-.93	.946	1.000 .375	.925	.976 .375
INAPPROPRIATE BUS ROUTE	AUTO BUS	.64-1.00	.908	.988 .333	.908	.988 .333
CARLESS	AUTO BUS	1.00-1.02	.759	.567 .868	.759	.567 .868

unity.

### Conclusions

Market segmentation techniques have been shown to be an effective procedure for including non-compensatory supply-side constraints in an attitudinal travel demand model. While it is infeasible to develop a supporting argument, it is contended that the choice of travel mode choice process modeled herein is sufficiently similar to many other consumer choice processes to warrant investigations of the effects of incorporating such constraints in models of these other choice processes.

Differences in structure of perception of choice alternatives were first found to be related to differences in supply-side conditions faced by individuals. Whereas these differences in perception appear to be mostly subtle, differences in choice behavior among groups of individuals faced with various supply-side constraints are marked. Treatment of supply-side constraints as constant conditional choice probabilities modifying preference probabilities led to minor improvements in choice model descriptive power. However the intuitively satisfying characteristics of these improvements for the various segments point toward fruitful future research.

It has been shown that significant aspects of the modal choice processes structure may be completely ignored by model estimations using a sample that is heterogeneous with respect to supply-side conditions. It is concluded that reliance on models based on total sample estimates can result in misleading, and often erroneous, forecasts of choice behavior.

### References

- G.H. Ball and D.J. Hall, "A Clustering Technique for Summarizing Multivariate Data," Behavioral Science, 12 (1967), 153-155.
- L.D. Burns, T.F. Golob and G.C. Nicolaidis, "A Theory of Urban Households' Automobile Ownership Decisions," Transportation Research Record, (in press). Available as General Motors Research Laboratories Publication 1785, (1975).
- J.G. Cragg, "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods," Discussion Paper No. 8, Department of Economics, University of British Columbia, (1968).
- D.H. Gensch and T.F. Golob, "Testing the Consistency of Attribute Meaning in Empirical Concept Testing," Journal of Marketing Research, 12 (1975), 348-354.
- T.F. Golob and R. Dobson, "Assessment of Preferences and Perceptions Toward Attributes of Transportation Alternatives," in Behavioral Demand Modeling and Valuation of Travel Time, Special Report 149, Transportation Research Board, National Academy of Sciences. Available as General Motors Research Laboratories Publication 1418, (1974).
- E.J. Gumbel, "Statistical Theory of Extreme Values and Some Practical Applications," National Bureau of Standards Applied Mathematics Series 20, No. 33, (1954).
- F. Hansen, "Consumer Choice Behavior: An Experimental Approach," Journal of Marketing Research, 6 (1969), 436-443.
- D.T. Hartgen, "Attitudinal and Situational Variables Influencing Urban Mode Choice: Some Empirical Findings," Transportation, 3 (1974), 377-392.
- D. McFadden, "The Revealed Preferences of a Government Bureaucracy," Unpublished Paper, Department of Economics, University of California, Berkeley, (1968).
- D. McFadden, "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, ed. Frontiers in Econometrics, New York: Academic Press, (1973).
- W.W. Recker and T.F. Golob, "An Attitudinal Modal Choice Model," General Motors Research Laboratories Publication GMR-1959, (1975).
- J.N. Sheth, "Attitudes and Other Competing Models of Consumer Choice Behavior," presented at Annual Convention of American Psychological Association, Chicago, August 30-September 3, (1975).
- H. Theil, Principles of Econometrics, (New York: John Wiley and Sons, 1971).
- W.L. Wilkie and E.A. Pessmier, "Issues in Marketing's Use of Multi-attribute Attitude Models," Journal of Marketing Research, 10 (1973), 428-441.

## A PSYCHOLOGICAL MODEL OF TRAVEL MODE SELECTION

Jagdish N. Sheth, University of Illinois

### Abstract

Based on a typology of the user expectations about alternative modes of travel, a psychological model is developed in this paper which hopefully provides an explanation for systematic choices people make about a specific mode for travel purposes including for commuting to work in an urban area. The user expectations are presumed to be five dimensional: functional, aesthetic-emotional, social-organizational, situational, and curiosity. The specific expectations about a mode on these five dimensions are determined partly by supply factors such as mode availability, mode design, mode operations and mode marketing; and partly by demand factors such as demographics, life styles, trip purpose, and prior awareness and satisfaction in the mind of the user.

### Introduction

The objective of this paper is to provide a comprehensive psychological model of travel mode choice behavior. Despite considerable research on people's mode preferences and choice behavior especially for urban travel, there is a conspicuous absence of any comprehensive, systematic theory of mode choice behavior anchored to the psychology of the user. This is largely due to the strong engineering and economic orientation to transportation research in which the search for the determinants of mode choice has been limited to either engineering designs and systems, operations and scheduling and economic costs associated with the mode usage; or to very broad socioeconomic-demographic characteristics of the individual users (See Vance, 1974 for an extensive and excellent bibliography on mode choice models).

Only recently there has been some empirical research on mode choice behavior which is based on the psychological processes of the users (Golob, Canty, Gustafson and Vitt, 1972). While this research has generated greater respectability for understanding the psychology of mode selection, there exists no comprehensive psychological model of travel mode choice behavior.

A psychological model of mode choice behavior is likely to be extremely useful in transportation research due to many diverse reasons. First, psychological factors are often the basis for deep-rooted habits and preferences of people which favor private transportation such as the automobile and which reject many good public transportation systems especially within an urban area. Second, we know relatively little about the psychological basis of public transportation alternatives because most research has been concentrated on the engineering-economic aspects related to the supply of these alternatives. For example, while we know quite a lot about the prestige and status symbols associated with the automobile, there is no comparable body of knowledge about the bus or the train. Third, psychological factors are often easier to bring about a change in the desired direction by managerial actions. Relative to engineering design changes, they are quicker, short-term strategies which are easier to implement with significantly less costs associated with them. Furthermore, often the costly engineering redesigns produce no desirable results because a particular mode is prefer-

red or disliked strictly due to psychological perceptual factors associated with that mode. In other words, a psychosomatic illness needs a psychosomatic and not a physical remedy. Finally, the enormous individual differences among user preferences across the cross-section of the population can only be understood properly by a detailed microlevel psychological analysis of the phenomenon. The traditional system-based variables such as schedules, fares and safety or even broad demographic variables such as income, education and age often do not account for a very large percentage of the variations in user preferences. Furthermore, from a public policy point of view, it seems only appropriate that the user psychology be fully understood to provide adequate consumer welfare and protection, and also to minimize the "catch-22" phenomenon so common in public regulation.

### Overview

The fundamental concept underlying the model of mode choice behavior is that each user has a five-dimensional subjective space of mode utility. Each mode is evaluated with respect to these five dimensions and all feasible modes are placed in the five dimensional hyperplane as points in the utility space. The five dimensional psychological utility space reflects the user's expectations with respect to the functional, aesthetic-emotional, social-organizational, situational and curiosity-based needs, wants, hopes, problems or barriers associated with travel behavior. The regular usage of a mode is determined by the user's evaluation of how well it performs on each of these five dimensions as compared to what he would like in a mode given his economic and time resource limitations.

There are several aspects of this model of mode choice behavior which need to be reviewed before fully describing it. First, it is a psychological model anchored in the subjective world of the user and not in the objective reality of mode characteristics. Therefore, the operationalization and testing of the model is, by definition, anchored at the individual user level. Aggregation at the total system level can be achieved only by assessing similarities and differences among individual users with respect to their subjective expectations. Often, this is likely to result in creating segments of system users who have similar perceptual mapping of the modes within each segment but contrasting mapping between segments.

Second, the model presumes that the impact of objective factors anchored to either mode engineering and operations or to user characteristics such as age, income and education on mode selection is not direct but mediated via the psychological expectations of the user. Thus, for example, the safety and scheduling of a public transportation mode such as the subway in a large metropolitan area will impact on a user's mode selection process only if they are mediated by his psychological expectations. Similarly, user's education and age may produce an impact on his mode selection if they determine or influence his psychological expectations.

Third, the psychological model of mode selection presented in this paper is capable of explaining and pre-



dicting only that mode choice which the user patronizes on an ongoing, regular and repetitive manner over a period of time. The model is not capable of predicting small, ad hoc deviations from regular mode usage which people occasionally make due to extraneous circumstances. For example, a person may regularly commute by private car but one day takes the bus to go to work because the car has broken down or is being serviced that day. Similarly, a person may usually take the train to go to work but one day goes in his personal car because he is going out of town straight from the place of work. In order, therefore, to predict each ad hoc use of a mode, we must go beyond the general psychological model and assess highly specific and unique situational factors. If the situational factors dominate in mode choice behavior, we would expect considerable degree of unstable behavior over time. In that case, the model developed in this paper is likely to be less useful. However, it is hoped that people manifest some regular patterns in their mode selection over time for each specific travel purpose.

Finally, the model is only a theoretical model which has not been as yet tested or validated. While past empirical research on mode selection provides some face validity to the model, it needs to be tested.

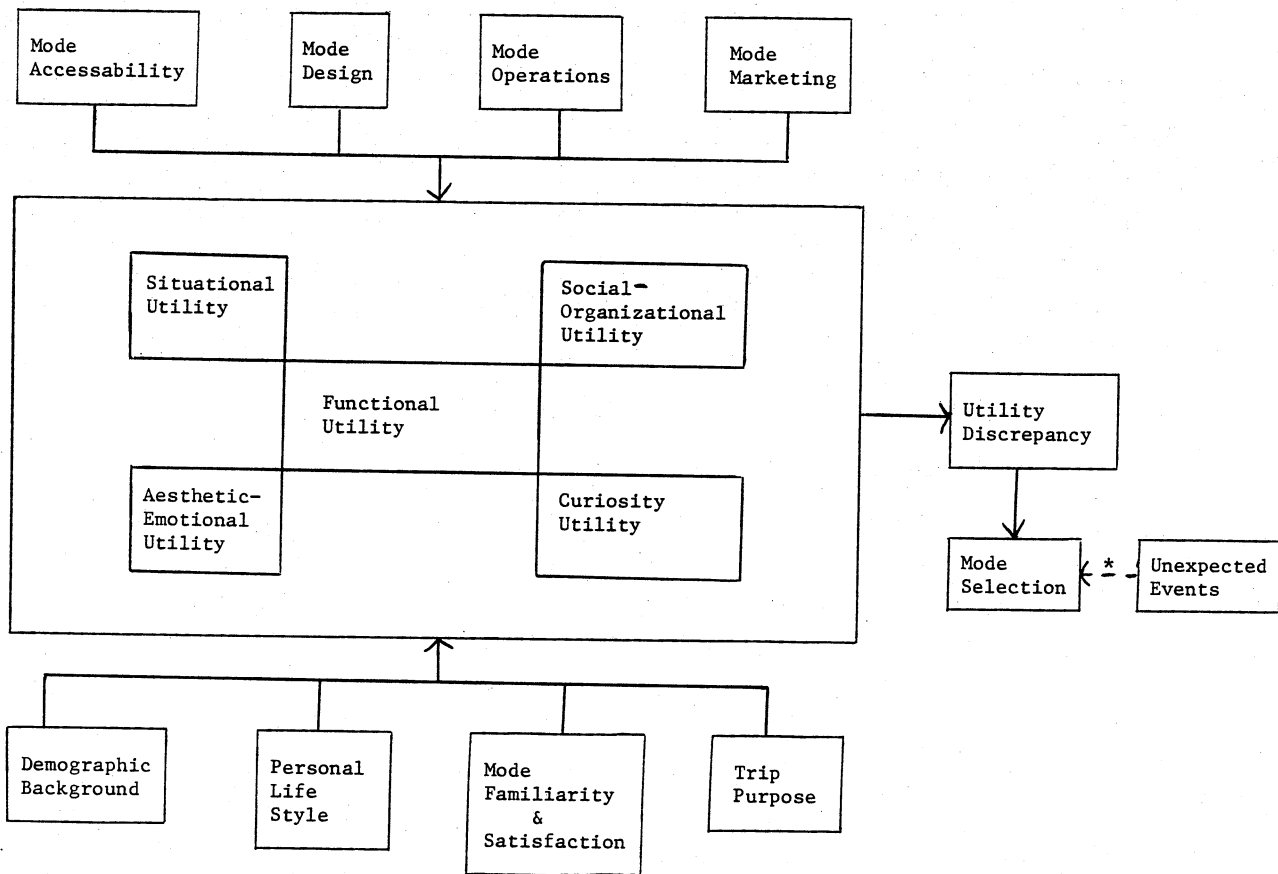
The model is represented in Figure 1. The rest of the paper describes various elements of the model in greater detail.

#### Dimensions of Psychological Utility

Based on a recent theory of individual choice behavior (Sheth, 1975), the psychological dimensions of mode utilities and underlying expectations are presumed to be as many as five dimensions. Not all the dimensions may be relevant to all types of travel behavior. In fact, the most important empirical research is likely to be the identification of dominant dimensions of psychological utility which vary from individual to individual, from one trip purpose to another trip purpose for the same individual, and from one mode to another mode across a cross-section of users. Each dimension of psychological utility associated with a mode will be described in detail.

1. Functional Utility. It represents the perceived utility of a mode which is strictly limited to it's performance as a mode of travel for a specific purpose. For example, the most common aspects of psychological functional utility related to various modes for commuting purposes tend to be arrival on time, direct route,

FIGURE 1. A PSYCHOLOGICAL MODEL OF MODE SELECTION



\* Needed to account for occasional deviations from regular mode usage.

safety, no transfers, having a seat, low maintenance costs, parking problems, etc. In other words, the psychological world of functional utility is limited to those mode expectations which directly relate to the function the mode is presumed to perform. As such, functional utility tends to more directly relate with the engineering design and system operations characteristics. Considerable research exists in transportation area on both the objective and the psychological assessments of a mode's functional characteristics (Vance, 1974; Golob, Dobson and Sheth, 1973). Most of this work has focused on the relative importance of functional attributes in determining people's mode selection and usage. Although perceived functional utility is clearly a function of a large number of mode attributes, this paper will not concentrate on how specific mode attributes combined together produce a certain level of functional utility or disutility in the mind of the individual user. It should, however, be pointed out that the functional utility is the net outcome of both positive and negative evaluations the user makes of a particular mode on a set of functional characteristics.

2. Aesthetic-Emotional Utility. The aesthetic-emotional utility is defined as that non-functional utility of a mode, which is anchored to fundamental values of the individual user in the emotive areas of fear, social concern, respect for quality of life, appreciation of fine arts, and other emotional feelings. The emotional-aesthetic utility toward a mode is often manifested in terms of style, interior and exterior decoration, comfort and luxury, as well as safety of the mode. However, often the user associates strong emotional feelings with a mode anchored to early experiences or some unexpected ad hoc experience. For example, a person may be involved in a train accident which may produce such a traumatic shock that the person is incapable of riding the train as it associates with very fundamental emotional feelings of death and survival. The aesthetic-emotive utility toward a mode is also developed by early childhood socialization processes. Thus, a person may reject the use of public transit based on a fear that he may catch diseases from other people's use of the same mode of travel.

The aesthetic-emotional expectations of a user toward modes of travel are often looked down by researchers as "irrational" motivations. However, often they determine his mode selection behavior and, therefore, must be fully researched.

3. Social-Organizational Utility. Travel modes especially in urban transportation tend to acquire certain stereotypes or imageries because they are selectively associated with certain socioeconomic, demographic and occupational groups of people. These socially anchored stereotypes produce certain mode utilities or disutilities which become an important dimension of user expectations. For example, carpooling or vanpooling may be looked down by some people if it is associated with lower income blue collar or clerical white collar workers. The suburban train may have social prestige since it is more used by upper socioeconomic professional people. Finally, certain bus routes within a city may acquire social disutility if it is primarily patronized by a select group of people such as Blacks.

Social-organizational utility is also nonfunctional in nature. Again, while considerable research on this type of perceived mode utility is available about the automobile, we know relatively little about public modes of transportation. The factors that tend to create social utility or disutility in a mode are the demographic factors of sex, race and age, the economic factors of income and price, and the social factors of

education and occupation.

4. Situational Utility. This refers to perceived utility of a mode which is strictly due to the circumstances surrounding its availability and accessibility. It also refers to the degree of utility or disutility a particular mode acquires as it relates to the total set of activities associated with a trip purpose. For example, while the automobile may be functionally a very good mode of travel for commuting purposes, the difficulty of parking, and the distance from place of work may produce disutility in it. Similarly, a person may go for shopping by bus because the spouse takes the single family car to work. People often prefer to drive on long-distance trips rather than fly due to the inconvenience of airline schedules of arrivals and departures as well as the time it takes to reach the airports.

Situational utility is similar to functional utility except that its presence or absence is often associated with the compatibility or difficulty with which related antecedent or subsequent activities involved in travel. The mode's own functional utility is, therefore, either enhanced or inhibited by the performance of related activities.

It should be kept in mind that a mode's situational utility is likely to vary from individual to individual. In fact, it is often determined by individual's own characteristics such as the area in which he lives, his household composition, and his personal life style.

5. Curiosity Utility. A mode may acquire additional utility beyond its functional utility simply because it is new or different. Many of the radical engineering innovations in transportation often generate curiosity utility which may temporarily increase the usage of the mode. This happens, for example, when new buses or train cars replace obsolete ones. Probably the new Bart system in the San Francisco Bay area may be, at present enjoying greater ridership because it is something new or different.

The curiosity utility is by definition, short-lived. Once a new innovation becomes old hat, it is more likely to be evaluated on the basis of other four types of customer expectations. However, when two modes of travel are equal in regard to other dimensions of psychological utility, the mode which offers innovative ideas, is likely to gain marginal utility due to curiosity, novelty expectations of consumers.

Each mode of travel is therefore, evaluated by a trip maker on each of the five dimensions of psychological utility. Based on the mode's utility vector, the trip maker decides on a particular mode as his regular mode of travel for that specific purpose such as commuting, shopping, vacations, visiting friends and relatives or sightseeing.

The model presumes that the individual, given his time and income constraints, has a desired profile of expectations with respect to what he seeks in travel behavior in terms of his own functional, aesthetic-emotional, social-organizational, situational and curiosity needs, wants, desires, problems and barriers. The discrepancy between the desired and perceived vectors of psychological utility will determine whether a particular mode will be acceptable to the individual. However, the utility discrepancy is presumed to be only one-tailed in nature. In other words, if a mode offers more than the desired utility on a specific dimension, the additional utility has no value to the individual. On the other hand, if it offers less than the desired utility on some other dimension, it will be regarded as

less than satisfactory. Depending upon the extent of one-tailed or positive discrepancy, a particular mode will become either acceptable or unacceptable to the individual. What is that critical level of positive discrepancy is an unknown parameter which can only be estimated by empirical research. However, it can be safely stated that the greater the positive discrepancy between desired and perceived utility of a mode, the less is the probability of it remaining acceptable to the user.

Mathematically, the Utility Discrepancy of a mode ( $UD_j$ ) is measured as follows:

$$UD_j = \sum_{k=1}^5 (U_{dk} - U_{jk}), \quad (1)$$

under the constraint that  $UD_{jk} = 0$  if  $U_{dk} < U_{jk}$ ; and where,

$UD_j$  = Utility Discrepancy of  $j$ th mode of travel for a specific purpose.

$U_{dk}$  = Desired utility on  $k$ th dimension

$U_{jk}$  = perceived utility of  $j$ th mode on  $k$ th dimension.

Finally, the choice of a mode as a regular mode of travel for a specific purpose is likely to be based on the principle of minimization of utility discrepancy. In other words, choose  $j$ th mode if  $UD_j$  is less than  $UD_1$ , any other mode of travel.

While the model may look highly rational from the decision-making process point of view, two things must be kept in mind. First, the basis for the mode choice contains at least four dimensions of psychological utility which are not based on the functional aspects of the mode performance. In other words, we presume that the consumer has a rational decision-making process even if what he desires in travel behavior may not be good for him. Secondly, as stated before, the model attempts to explain only those mode selections for a specific purpose which are stable over time except for occasional, ad hoc changes. The model is not capable of predicting a person's mode choice for a specific day and for a single trip. Such specific predictions are more difficult and probably not amenable to model building effort because too many unpredictable factors tend to be involved in those choices.

#### Determinants of Psychological Utility

The vector of psychological utility is likely to vary from mode to mode, from individual to individual and from trip purpose to trip purpose. Given this three-way variability in the psychological utility, it becomes necessary to isolate factors which seem to determine it. It would appear that some of the determinant factors are likely to be supply-oriented and others are likely to be demand-oriented. We will discuss each type of determinant factors below.

1. Supply-Oriented Determinants. How a person will perceive the psychological utility of a mode of travel is at least partly determined by what the supplier of the mode has to offer to the customers and the manner in which he offers mode selection relative to other suppliers of other modes of travel. The psychological utility of a mode is presumed to be partly determined by at least four supplier-oriented factors. They are mode availability, mode design, mode operations and mode marketing.

Mode availability includes ease of accessibility of a mode to the trip makers. It also includes the total network of the mode system. For example, in a bus sys-

tem, it is important to consider the total routing system, distance from places of work, residence, shopping and recreation as well as the frequency with which it is accessible to the trip makers. Similarly, the availability of a car implies a relatively large amount of capital expenditure and its accessibility for a trip purpose depends on the highway system. In general, it is safe to state that the private modes of travel such as the automobile have a far better availability factor in their favor relative to most public modes. This has resulted in greater functional and situational utility in the automobile as a mode of travel in general.

Mode design is a second supply-oriented factor. It includes the variety of product or service as it is offered to the customers. Thus, in the case of the automobile, it includes variety of models, styles and conveniences as well as distinct features relevant to its performance as a mode of travel. Similarly, the subway cars, their interior design, seating arrangements, and station conveniences constitute mode design of a subway system. Mode design is not limited to the physical carrier but also includes all related facilities. For example, parking lots and parking spaces, the highway design and scenery will be relevant elements of automobile product design. Similarly, the stations, platforms, newsstands and restrooms will be part of the subway mode design. The mode design variable provides both the functional and aesthetic-emotional utilities to a mode. Furthermore, it often is capable of injecting curiosity utility by planned changes in secondary or nonfunctional characteristics. Examples include styling changes in automobiles, buses and subway cars.

Mode operations refers to the actual usage related man-machine interface involved in the daily operations of a travel mode. For example, in a bus system, it includes the driver, the procedure for paying the fare, and the procedures for getting in and out of the bus at predetermined stops. Similarly, the mode operations in carpool include the rotation of carpoolers as drivers, the specific route taken to pick up and drop off carpoolers as well as fuel and maintenance aspects of the car. It is often the mode operations which either enhances or inhibits the functional and situational utility of a mode. It is not at all uncommon to find the same system put under a new management which actually has saved the mode of travel from going bankrupt. The managerial talents are extremely useful or important in this regard. Amtrak is a good example of this supply-oriented factor.

Finally, mode marketing is an important determinant of the psychological utility of a particular mode. Although marketing is somewhat narrowly defined in this paper, it includes both price and promotion strategies. The role of mode marketing is primarily to enhance the functional, situational and social utilities involved in a mode. It can also perform the role of providing adequate information to the trip maker.

The supply-oriented factors combined together generate differential psychological utilities for different modes of travel. It is these factors which often create mass acceptance or rejection of a mode in the market place.

2. Demand-Oriented Factors. It is not sufficient to presume that the psychological utility is fully determined by the supply-oriented factors. If that were the case, the task of management and public policy would be relatively simple. The psychological utility of a mode is also determined by a set of demand-oriented factors. These factors tend to produce differential psychological utilities for the same mode of travel among a cross-section of users resulting in the same

mode accepted by one group of customers and rejected by others.

There are four distinct demand-oriented factors. They consist of (1) personal demographics of the trip maker, (2) his personal life style, (3) his familiarity and satisfaction with the mode and (4) purpose of his trip.

Demographics of the trip maker consists of sex, race, age, income, education and occupation of the trip maker. In general, it includes the individual's life cycle and socioeconomic status. There is no question that personal demographics especially sex, age and race heavily influence what the individual desires by way of psychological utility in a mode. In addition, income is likely to influence these desires. Considerable research exists today to support the influence of demographics on people's mode preferences (Hille and Martin, 1967; Wynn and Levinson, 1967; Bostick and Todd, 1966; Golob, Cauty, Gustafson and Vitt, 1972).

Life style refers to an individual's allocation of time and money in the daily activities of one's life including specific choices he makes in terms of food, shelter, clothing, recreation, religion, work, community involvement and the like. Life styles tend to reflect individual's fundamental value system. Often they influence a person's psychological utility in a manner which is independent of demographic factors. Recently, considerable research has been generated in consumer research to indicate that life styles tend to provide insights into consumption differences among otherwise homogeneous demographic segments (Wells, 1974 and 1975). Very little research in transportation area has been undertaken so far to measure the impact of life styles on mode choice behavior either directly or mediated via psychological expectations and preferences.

Mode familiarity and satisfaction is an obvious but very useful determinant of mode expectations. As has been pointed out by Bostick and Todd (1966), people tend to choose a given mode and stay with it if they are satisfied even though other modes may be more appropriate for them. Often, an individual does not choose a mode of travel simply because he is not even aware that it exists, or knows so very little about it as to ignore it. Similarly, people who tend to use a mode regularly form habits which are often difficult to change.

The final demand-oriented factor is trip purpose. This is included primarily to account for differential preferences of a mode on the part of the same individual for different trip purposes. For example, many people take the train or the bus to go to work but prefer to go in a personal automobile for shopping even if the place of work and shopping are in the same location. Each trip purpose is presumed to have somewhat different set of needs, wants, desires and problems associated with it so that the same mode may be best for one purpose but quite inferior for another. This is dramatically true when the businessman drives to work but flies for out-of-town business activities.

The influence of the supply-and demand-oriented factors on the vector of psychological utility is presumed to be monotonic and additive. Mathematically, we can state the formal relationship as follows:

$$U_{ijk} = \sum_{l=1}^r a_{jkl} S_{ijkl} + \sum_{m=1}^p b_{jkm} D_{ijkm} \quad (2)$$

where  $U_{ijk}$  = Utility vector of  $i$ th individual for  $j$ th mode for  $k$ th trip purpose

$S_{ijkl}$  = Evaluation of the  $j$ th mode for  $k$ th trip purpose by  $i$ th individual on  $l$ th supply-

related factor.

$D_{ijkl}$  = measurement of individual  $i$ 's  $l$ th demand-related factor with respect to  $j$ th mode for  $k$ th trip purpose.

$a_{jkl}$  and  $b_{jkl}$  are constants to be estimated.

It should be noted that some of the variables will have less than four subscripts if they are generalized measures affecting all modes or all trip purposes. These include the demographic and the life style variables as well as some of the supply-oriented variables.

#### Conclusion

A psychological model of mode choice behavior has been presented in this paper which has the fundamental basis of evaluation of a mode on a five dimensional utility analysis. It is suggested that the mode selected by the individual for regular usage, for a specific purpose is likely to be based on the minimization of discrepancy between what is desired and what is perceived of a mode in regard to the five dimensions of psychological utility.

A number of supply and demand-oriented factors are suggested as determinants of the psychological utility of a specific mode  $j$ , for a specific trip purpose  $k$  in the mind of individual  $i$ . These include mode availability, mode design, mode operations and mode marketing on the supply side, and demographics, life styles, prior familiarity and satisfaction with the mode as well as trip purpose on the demand side.

It is hoped that the model is comprehensive enough to understand the phenomenon of travel mode behavior. Whether the model will be validated or not remains to be tested.

#### References

- T. A. Bostick and T. R. Todd, "Travel Characteristics of Persons Living in Larger Cities," Highway Research Record, No. 106, (Washington, Highway Research Board, 1966), 52-57.
- T. F. Golob, et. al. "An Analysis of Consumer Preferences for a Public Transportation System," Transportation Research, 6, 1972, 81-102.
- T. F. Golob, R. Dobson and J. N. Sheth, "Perceived Attribute Importance in Public and Private Transportation," in H. W. Hopfe (ed.) Advancing, Applying and Teaching the Decision Sciences, (Atlanta: AIDS, 1973), 7-10.
- S. J. Hille and T. K. Martin, "Consumer Preference in Transportation," Highway Research Record, No. 197 (Washington: Highway Research Board, 1967), 36-43.
- J. N. Sheth, "Toward a Model of Individual Choice Behavior," paper presented at ESOMAR Seminar on Market Modeling, June 11, 1975.
- M. Vance (ed.), Modal Choice Models: Introduction and Selected Bibliography (Monticello, Ill.: Council of Planning Libraries, 1974).
- F. H. Wynn and H. Levinson, "Some Considerations in Appraising Bus Transit Potentials," Highway Research Record, (Washington, Highway Research Board), No. 197, 1967, 1-24.
- W. D. Wells, Psychographics and Life Styles (Chicago: American Marketing Association, 1974).

W. D. Wells, "Psychographics: A Critical Review,"  
Journal of Marketing Research, 12(May, 1975),196-213.

EXPLORATORY GROUP INTERVIEW IN CONSUMER RESEARCH: A CASE EXAMPLE<sup>1</sup>

Thomas D. Dupont,<sup>2</sup> Oxtoby-Smith Inc.

Abstract

Several applications of focus group interviewing are described, together with the unique advantages of this technique. The paper focuses on the use of exploratory group interviewing as a preliminary to a larger, quantified study. A case example of research conducted for Volvo of America Corporation is presented as an illustration of the functions of exploratory group interviewing.

Introduction

"Look, my budget's tight and we don't have much time. Let's just do a couple of groups."

"We need to talk with these people to find out what they think. We've got to really dig in below surface opinions. So let's round some up and do some groups."

"I understand the client's problem, but I don't even know who we should survey or what questions we should ask. Let's recommend that we start with a couple of groups and design the quantification after we have a better idea of what the relevant issues are."

Sound familiar?

I'm not saying that those are the only reasons why market researchers use focus group interviews, but my guess is that those three scenarios account for a substantial majority of them.

Is there anything wrong with this?

Certainly not. Those three situations are made to order for group interviews. In fact, there are a variety of very good reasons why one would wish to conduct group interviews:

1. Group interviews can be conducted very quickly, and relatively inexpensively.
2. They sometimes provide the only mechanism by which a senior researcher can talk face-to-face with consumers. Individual depth interviews, the alternative, are simply too expensive in most circumstances.
3. Group interviews provide the opportunity for the non-research client -- the Product Manager, Vice President - Marketing, Sales Manager -- to listen to what consumers are saying about his product and the way he sells it. What he hears is sometimes painful, but almost always enlightening.

<sup>1</sup>The author expresses his appreciation to Lars Samuelson, Vice President -- Marketing, Volvo of America Corporation, for permission to publish this paper, and for supporting the research described herein.

<sup>2</sup>Thomas D. Dupont is Vice President, Oxtoby-Smith Inc., New York, N. Y.

4. Group interviews are flexible. The moderator is not tied to a fixed sequence of questions, but can skip topics which seem to be unproductive and zero-in on areas which are productive. Sometimes, the ability of a good moderator to exploit such "targets of opportunity" can result in purely serendipitous findings of great importance to the client. Such results emerge less frequently from field interviews.

5. Participants in the group can interact and stimulate one another. While a group interview is in no sense similar to a group therapy session, and the moderator is in no sense a therapist, the plain fact is that the group situation -- provided it is a secure and comfortable one -- often encourages respondents to disclose attitudes and behaviors which they might not admit in an individual interview situation.

Just as there are a number of good reasons for conducting group interviews, there are a number of problems which can be addressed in group interviews. By way of example, within the past year, Oxtoby-Smith has used group interviews:

1. To explore consumer reaction to new product concepts;
2. To explore consumer response to both advertising concepts and finished ads;
3. To generate ideas for new products;
4. To explore consumer response to package designs and labeling; and
5. To explore differences in perception between a live demonstration of a product and a filmed presentation.

However, the most important function of group interviewing at Oxtoby-Smith is as a device to guide the design and conduct of a subsequent large-scale quantitative survey. In this context, group interviews are used:

1. To identify and understand consumer language as it relates to the product category in question. What terms do they use? What do they mean?
2. To identify the range of consumer concerns. How much variability is there among consumers in how they view the product and in the considerations which lead them to accept or reject the product?
3. To identify the complexity of consumer concerns. Are there a few simple attitudes which govern consumer reaction toward the product, or is the structure complex, involving many contingencies?
4. To identify specific methodological or logistical problems which are likely to affect either the cost of the subsequent research, or our ability to generate meaningful, actionable findings at all.

I'd like to describe a project we recently completed which illustrates these four functions of exploratory group sessions.

## A Case Study

### Background Of The Case

For the past six years, Oxtoby-Smith has been conducting consumer research for Volvo of America Corporation. During that period, we have witnessed a substantial change in the marketplace as it affects Volvo:

1. Consumers seem increasingly interested in cars like Volvo -- functional, durable, compact and safe.
2. Until very recently, Volvo and a few other imports had this segment to themselves; prior to Ford's introduction of the Granada there was little direct domestic competition in this segment.
3. Most importantly, because of successive devaluations of the dollar and other economic factors, the price of the least expensive Volvo has risen from about \$3,300 in 1972 to about \$5,500 in 1975.

It became increasingly apparent to Volvo marketing management, and to Volvo's advertising agency, that whereas Volvo used to compete primarily with a small group of imports (mainly Audi), in the future Volvo, because of its price level, would be increasingly competing with domestic luxury cars (Buick Electra, Chrysler Cordoba, Ford Elite, and even Cadillac and Lincoln Continental).

In consequence of these changes, we decided to undertake a study to explore the differences between Volvo buyers and Volvo considerers (people who thought about buying a Volvo, but in the end bought some other car). We wanted to learn the reasons for buying a Volvo, and the resistances to buying a Volvo.

### The Study Design

Our study design, as you might guess, involved inquiry among two respondent populations -- people who had recently bought Volvos, and people who had recently considered buying Volvos, but instead bought competitive makes.

Finding buyers is easy -- Volvo has lists. Finding considerers is not so easy, especially since a previous attempt at getting Volvo dealers to record the names and addresses of all prospects was a dismal failure.

Accordingly, we decided to buy lists of buyers of new car models who we felt would be most likely to have considered Volvo. This selection of models was based upon the premise that the cars Volvo buyers reject are the same cars which Volvo considerers buy. Since we knew from previous research what other cars Volvo buyers consider, we had our list of models. The design called for two phases of research:

1. An exploratory stage, involving four group interviews -- two with Volvo buyers and two with Volvo considerers (prior research had taught us that buyers and considerers don't mix well; buyers dominate the sessions by proselytizing about Volvo).
2. A quantification phase involving telephone interviews with 400 first-time Volvo buyers and 200 Volvo considerers.

In the remainder of this paper, we'd like to describe what we learned in the exploratory group sessions, and how what we learned affected the design and conduct of the quantification study.

## What We Learned From The Sessions Themselves

We had conducted a great deal of research for Volvo in the past and had generated a number of hypotheses about what makes a considerer become a Volvo buyer. In spite of this extensive background in the product area, however, we still learned important things in the groups which permitted us to conduct a quantification study which was both more insightful and less expensive than we would otherwise have conducted.

Let me give you a few examples.

Volvo considerers differ sharply in the way in which they consider Volvo. One of the most important things we learned -- which confirmed what we suspected -- was that there are a number of different ways in which considerers considered Volvo. Some considered Volvo very seriously, and narrowed the choice of cars to Volvo and one other make. Others considered Volvo seriously, but it was not among the cars which survived until the final decision. (In selecting respondents for the groups, we asked whether they seriously considered a Volvo, and only serious considerers were recruited. Prior experience had taught us that a question like, "Did you consider buying a \_\_\_\_\_" is a very weak question. Many consumers will say they considered buying a product even if that consideration was very fleeting and casual.) Even within the "serious considerers," the ways of considering Volvo varied enormously:

1. Some test drove the car and haggled over price with the dealer.
2. Others evaluated the car carefully, but did not get to the point of negotiations with the dealer.
3. Some considered the car seriously without ever visiting a showroom, relying instead on word-of-mouth and ratings of the car in magazines like "Consumers Report," "Motor Trend," and the like.

Accordingly, we felt it crucial in our quantification study to include questions which would enable us to segment considerers according to the degree to which they seriously considered Volvo and the methods they used to arrive at their purchase choice. To that end, we designed a question to determine which specific actions were taken in considering Volvo, such as talking with Volvo owners, reading evaluative articles in magazines, paying more attention to Volvo advertising, visiting a Volvo showroom, negotiating over price, and others.

Volvo buyers and considerers are highly segmented. Since Volvo's share of the U.S. auto market is less than 1%, it would be reasonable to expect that Volvo buyers, and to a lesser extent, Volvo considerers, are relatively homogeneous in terms of demography, attitude, and preference. However, while it is true that these groups are more homogeneous than U.S. car buyers at large, it is nevertheless the case that segmentation is very real within Volvo's market.

1. Volvo buyers range from those who bought Volvo as a less expensive alternative to Mercedes to middle-income Americans who must strain to afford the car. They range from first-time import buyers to those who would never consider a domestic car.
2. Volvo considerers vary even more widely than do Volvo buyers, from "typical domestic buyers" interested in styling, power, status, and a familiar name to "typical import buyers" interested in economy, functionalism, durability, and "foreign craftsmanship."

These group interviews strongly suggested that the primary variable differentiating the Volvo buyer from the Volvo considerer was the set of concerns that individual brought with him to the car-buying process. Further, it was readily apparent from the groups that these concerns varied widely, implying that it would be necessary in the quantification to measure the importance to the consumer of a number of characteristics of a car, such as safety, exterior styling, anticipated cost of service, and so forth. In consequence, we prepared a list of 26 automobile attributes -- some drawn from prior research and some new ones based upon the group session findings to present to consumers in the quantification. It was our hope (subsequently realized) that importance ratings of these items would prove to be extremely powerful in discriminating between Volvo buyers and Volvo considerers.

However, the desirability of this approach left us with a methodological dilemma. It was important that we not only learn the importance of these 26 factors to Volvo buyers and considerers, but also that we learn, among considerers, how Volvo stacked up, on the factors they considered important, against the car actually bought.

The logical alternative, after having the respondent rate all 26 factors for importance was to have him rank the five factors which played the largest role in the automobile selection decision, and then rate Volvo versus the competitive car on those top five factors. Clearly, it would be expecting too much to have the respondent perform this task in a telephone interview. It was critical, we felt, that the respondent have the 26 factors in front of him while he was going through the rating process. Thus, we were left with two alternative methodologies -- personal interviews or mail. As we will see, the screening experience in recruiting respondents for the sessions strongly influenced our choice between these two alternatives.

#### What We Learned From The Recruiting Process

As was mentioned earlier, prospective participants for the group sessions were recruited from lists of recent buyers of cars felt to be competitive with Volvo. In order to be invited to a session, the respondent had to indicate that he seriously considered buying a Volvo.

As it happens, the results of the screening for serious considerers proved very instructive; we had to contact about 1,000 recent new car buyers to recruit approximately 20 focus group respondents. While we knew we could expect a higher cooperation rate in the quantification, since respondents would not have to leave their homes, the fact remained that the incidence of Volvo considerers, even among buyers of selected models, was very small. Accordingly, a quantification study conducted via telephone interviewing, to say nothing of personal interviewing, would substantially exceed the budget which Volvo had set aside for this project.

In that circumstance, we decided to conduct the study as a mail survey -- a decision which permitted us to collect the rating information which we could not collect over the telephone, while at the same time saving Volvo a considerable sum of money (we estimate about \$10,000). The mail survey went smoothly, the quantification verified some of the hypotheses we had generated in the group sessions (and refuted others), and the results of the study have since played an important role in the development of Volvo advertising and marketing strategies for 1976.

In summary, this example illustrates quite well the value of conducting a few exploratory group interviews prior to a larger, quantified study. Such an exploration will not always pay for itself in a more efficient final research design, as this one more than did, but it will nearly always lead to a subsequent survey richer in content, more adroit in interpretation, and more actionable to marketing management. That is why Oxtoby-Smith -- as well as many of our competitors -- makes a standard practice of conducting some exploratory group interviews prior to any extensive data collection effort, even when the product area is one with which we are highly familiar.



## PREPARING FOR GROUP INTERVIEWS

Melanie S. Payne, Elrick and Lavidge, Inc.

### Abstract

Proper preparation for group interviews is essential to achieve meaningful study results. Responsibility for conducting the groups should be given to a well-trained qualitative researcher whose job it is to understand the research objectives and how the results will be used. Other aspects of preparation include writing the discussion guide and supervising recruitment of respondents. Prior to the groups, decisions must also be made concerning where the sessions will be held and the type of audio or video recording that will be done.

### Preparing for Group Interviews

We have just heard from Al Goldman about the development of interviews and the way they are used, and shortly Myril Axelrod will discuss the procedures that should be followed when conducting groups. Once you have decided that groups are the proper research tool to use and before you actually run the sessions, considerable attention should be devoted to preparing for them. Because group interviewing appears to be such an easy thing to do and the process itself seems so casual, the important step of preparation is often neglected and left to chance.

One of the most crucial elements of preparing a group study is to place the responsibility for it in the hands of a well-trained moderator/analyst. This may seem obvious, but because the technique has gained such popularity and because it seems so simple, more people are getting into the act, and they are often the wrong people. I have seen too many instances where, to save money, a company will turn the job of conducting groups over to a secretary. They give the assignment to her because she is a pleasant individual who gets along well with people and she is a good conversationalist. That is not enough, however, because what is required is someone thoroughly grounded in the theory and technique of what group interviewing is all about.

Once the assignment has been placed in the hands of the right person, the first order of business in preparing for a group interview is to gain a thorough understanding of the problem. Occasionally this means helping the client understand the problem, too. Many clients, of course, know precisely what their needs are and how they will use the results, but all too often the person requesting the focus group research does not really know what he wants to find out or what he will do with the information once the study is completed. In this latter instance, he may know simply that he has a problem, and because the group approach is simplistic, he decides that is what he will do, without really thinking through the issues involved.

In either case, it becomes the researcher's job to clearly spell out the reasons the research is being done, the specific areas to be covered in the groups, and how the results are going to be used once the research is completed. Let me give you an example. Assume you have been asked to do some work for a company that has, (1), done virtually no consumer research and, (2), does not have an understanding of what its customer's attitudes about the company's products are. In this instance, your first job will be to educate the client about what the group interview technique can and

cannot do. Secondly, when you actually conduct the group discussions you will probably have to devote some of your efforts to gathering basic attitudinal information about the product category in general and your client's product in particular.

The approach just outlined will be very different from one you would use with a client who has a long history of researching his products and now wants some group sessions to help him develop the concept for a line extension of an existing brand. In these sessions you would shorten the background gathering and move quickly to the heart of the matter -- the new product. My point here is that all groups do not follow the same format and the correct approach to use is a function of the nature of the problem to be tackled.

As you gain an understanding of the problem, you should begin making a list of specific questions to be explored in the research. Some of these questions will be raised by the client, others will occur to you, and the combined list will become the skeleton for your discussion outline. The easiest and most useful way to handle the questions is to organize them by areas or topics so they become clustered in logical groupings. For example, if you are doing some research on a new hot breakfast cereal, you might want to begin with a general discussion of what is served at breakfast and then refine that by having respondents describe weekday versus weekend breakfasts. From there you might move on to focus on cereals, both hot and cold, then narrow down to hot cereals only. At this point you would probably introduce the concept for the new product, and you might even have some actual product prepared for respondents to taste. Just like a good story, a group discussion should have a beginning, a middle, and an end. Typically, as in this example, the movement is from the general to the specific.

The discussion guide itself should be fairly simple -- two or three typed pages, double spaced. It is an outline of the key areas and questions to be covered and not an exhaustive dissertation of every conceivable issue that might be probed for. If you understand what you want to do in a group and have the questions pretty well fixed in your mind, you should hardly have to refer to your outline once you are in the group setting itself. I suppose we all have our own little tricks for achieving this goal. My own personal quirk is that I must sit down at the typewriter and personally bat out the discussion guide I am going to use. Even if someone has prepared a perfectly adequate guide for me, I cannot seem to internalize it unless I type it myself.

Another essential element in the preparation of a guide, whether you have my kind of hang-up or not, is an adequate amount of time to think about the topic of your discussion. I feel that a minimum of a couple days is necessary to allow the topic area to sufficiently sink in. I don't mean to say that one should spend two or three days working on an outline, but you do need time to mull over in your mind what you are going to cover in a session. I find that I do this when I am preparing dinner, or riding the train to work, putting on my make-up, or whatever. No moderator, no matter how experienced he or she is, can be handed a group discussion guide half an hour before a group and be expected to conduct a first-rate interview.

At the same time you are involved in the process of understanding the problem and preparing the discussion guide, recruiting for the group should be getting under way. This will generally be taken care of by someone other than yourself. Just as the moderating should not be handled by an amateur, recruiting too should be supervised by a specialist. It may seem simple enough to gather together several "warm bodies" for a group discussion, but once again I have seen the amateurs who thought this was such a snap fall flat on their faces. So the moral of the story is pay a little more and find someone who knows what they are about to recruit your groups.

Your responsibility in working with the recruiters is to tell them exactly who you want and who you don't want in your groups. Let's look at the exclusions first. You do not want people who work for marketing research companies or advertising agencies or those employed by your client or any of his competitors. This is fairly obvious. You will also want to screen out certain groups who, for some reason, might bias your results. One of the worst examples I ever heard along those lines happened to a researcher who was doing some conceptual work on baby food and wanted to interview mothers of infants concerning feeding practices. The project director did not learn until the group had assembled that all the respondents belonged to the Leche League. These are women who breast-feed their babies and, not only do they breast-feed, they belong to a league of breast feeders. Needless to say, their responses were hardly typical of all young mothers.

The recruiter should have known better than to have pulled a stunt like that, but that little episode which happened more than 10 years ago taught me to never leave anything to chance. You should always stipulate that you want respondents recruited from a wide area, and that you do not want them to come from the same neighborhood or church or club or ethnic group. My own feeling is that it's okay if two people, but no more than two, know each other in a group, that is they may come in pairs if necessary. This is simply being realistic because often a woman will not agree to participate in a session unless she can come with a friend.

The question of professional respondents always arises -- these are the people who make a habit or perhaps even a living of participating in groups. The opposite of the professional respondent is the so-called virgin respondent -- the person who has never taken part in a group session. What you should ask for and can realistically expect lies somewhere between these two extremes. If a recruiter is providing you with the same faces again and again, you had better take your business elsewhere. Keep in mind that if you have seen those people repeatedly you can be pretty sure that they have recently been respondents in someone else's groups, too.

On the other hand, be fair with your recruiter and do not demand that they provide you with virgin respondents. There is simply too much group interviewing being done these days for that to be a reasonable request. I once calculated that with the number of groups being done each week in Chicago, if everyone demanded virgin respondents we would have exhausted the entire population of the city in a period of eight years. I generally ask that respondents in my sessions have not participated in a group within the preceding six months or perhaps within the preceding year. You will find, however, that even with this sort of screening, people enjoy participating in groups so much that they will lie about not having been interviewed so that they can come back soon again.

After you have stipulated whom you don't want attending, you must be very specific about whom you do want. For example, if you are conducting a study on a new brand of frozen french fries, you do not simply ask for people who eat french fries. Rather, you ask for women who have bought, and prepared for their families, at least two pounds of frozen french fries within the past month. Set your qualifications specifically and precisely and you won't be disappointed by coming face to face with a group of women who have little interest in your product category and who, therefore, cannot be the least bit of help to you.

How many respondents to include in a group is another key issue and one that I find myself getting into arguments about all the time. I am going to be very dogmatic on this point and say that no group discussion should ever have any more than eight respondents. If you, as a moderator, are performing your job correctly, six or seven people is perfectly adequate. I am appalled at the trend which seems to be in vogue now to demand ten or twelve respondents per group. What we want to accomplish with these people we have recruited is to carry out a discussion in which everyone participates as a group. With more than eight people, this process simply cannot occur naturally. The group breaks down and you, as moderator, are faced with the chaos of two or more splinter conversations going on simultaneously. To maintain control you then have the choice of policing the group conversation, thereby destroying the dynamic process you wanted to set up in the first place, or you retreat to the position of having to conduct a series of individual interviews in a group setting. Either solution is unsatisfactory.

At the time respondents are recruited for a group they should be told how long it will last. I generally ask people to be prepared to stay two hours even if I don't expect the session to last that long. What you want to avoid is underestimating the time involved. Men and women who come expecting to spend an hour get very fidgety when that hour is up. At that point they will say anything just to get the group over with and get out, leaving you with a lot of results that you can't have very much faith in. Two hours is about the outside limit for a productive group session. Occasionally some respondents get so wound up with a topic they will, of their own choosing, stay on and on. But normally a group will have said all that it's going to say by the end of the second hour.

It is possible to conduct a group interview literally anywhere there is enough room to seat the respondents and yourself. I have done them in church basements and in posh boardrooms and just about anywhere else you can imagine in between. I personally do not feel that the setting is nearly as important as the tone you set and the rapport you establish with the respondents. If they feel you are on the same wave length with them, you'll get a good interview no matter where you are.

The most desirable facility these days, especially from a client's point of view, is an interviewing room equipped with a one-way mirror and a comfortable room from which to view the group in progress. That viewing room should be as soundproof as possible so that the conversations of the observers cannot be heard by the respondents. Some research companies seem to be especially proud of the fact that they have well-stocked open bars in their viewing rooms. I find this not only unnecessary but offensive because it tends to turn the group session into some sort of charade rather than the serious business that it should be.

The furnishings in the interviewing room, per se, should be comfortable, but they need not be elaborate.

There are some who prefer a living room atmosphere with coffee tables and easy chairs. My own preference is for a large conference room table that will seat eight or nine people without crowding. I prefer the table because I think it gives respondents something to hang on to -- both literally and figuratively. Many men and women are very nervous when they come to a group discussion because they don't know what to expect. If they can sit down at a table, set down their coffee cups and fill out a short questionnaire covering product usage and demographics, it seems to put them at ease. Theoretically, people can accomplish the same end sitting in easy chairs clustered around coffee tables. But I find that in such a setting there is a lot of fumbling with clipboards, pencils, coffee cups and ashtrays. Also, women feel compelled to pull their skirts down to cover their knees, especially in a room with a very obvious one-way mirror. I prefer to avoid this altogether by conducting sessions around a conference table.

The presence of such a table also sets a somewhat businesslike tone to the session, and I don't think that's at all undesirable. While you want the respondents to be comfortable, this is a research session and not a neighborhood coffee klatch, and they surely know that ahead of time. They are not there just to shoot the breeze; they are being paid for their participation, and their conversations are being tape recorded. So it seems to me that to force all of this into a simulated home setting doesn't really accomplish very much.

I mentioned tape recording, and the extent to which you will record the proceedings is another decision you will have to make. The less gadgetry you have to worry about, the easier your life will be. If the facility you are using is equipped with an overhead microphone and a sound system, this can be an advantage. But if you don't have this, it's no cause for alarm. Any reel-to-reel tape recorder that can pick up the conversation will do. Also, I routinely take along a small cassette tape recorder to use as a back up. These recorders generally run on batteries, which can be a lifesaver if the electricity fails -- something that has happened to me more than once.

In at least 99 percent of the group interviews you do, audio taping is all you should need. But I have found as this group interviewing business becomes more elaborate, clients are requesting video taping more frequently. My guess is that in many of these cases the video tapes end up being stored on a shelf and are never looked at again. Few people have the interest or stamina to sit and stare at eight to ten hours of film after the fact, so if the tape is going to be used at all it should be edited. This is an extremely time-consuming process, and no one can have an appreciation of how much time it takes until he has been through it once. And believe me, once is enough. Both the filming and the editing are expensive, and my feeling is that much of the time and money spent on video tape might have been better invested elsewhere. But it is the client's money, and he can spend it in any way he chooses.

Just remember that if you are planning to videotape you must meet some special conditions, including lots of ceiling light and no windows or back lighting behind the respondents (or you will end up with shadows instead of faces). Also, to obtain a tape that has any visual interest at all, you should hire a professional cameraman whose job it is to follow the discussion with the camera and zoom in and out to pick up the person speaking. The alternative to this is a fixed camera with a wide-angle lens that is visually no better than an audio tape.

In the past 15 minutes I have taken you through approximately two weeks worth of preparation for a group interview. Some of the issues I covered may have seemed obvious, others trivial, but from where I sit, all are essential. Going back to the point I made at the outset, to the uninitiated, the amateur, or the outsider, group interviewing appears to be a simple, casual process that one just waltzes through. But as the case with everything from raising children to making a souffle, if it ends up being done well, it didn't just happen; it was planned for.

## THE DYNAMICS OF THE GROUP INTERVIEW

Myril D. Axelrod

Young & Rubicam International, Inc.

Because I am employed by one of the world's largest advertising agencies and work closely with a variety of clients, including many of the major marketers, I have had the opportunity to witness, in recent years, a growing feeling of insecurity with focussed group research and a growing desire to have some kind of yard-stick by which to "measure" the kind of focussed group services that are being bought. At least three of the client "giants" I know of have even gone so far as to issue virtual bans on the use of this kind of research unless it can meet rigid standards of positioning and quality.

Part of the concern about focussed groups, my client sources report, stems from the very serious problem of misuse of the observations drawn from them, and the mistaken impression that such research can be used to circumvent or substitute for larger-scale quantitative exploration or to provide definitive direction in areas where it is not able to do so.

The question of when to use focussed groups and the very important contributions that good focussed group research can and does consistently make will be discussed by other panelists. I won't, therefore, dwell on it, except to express my hope that we, as researchers, will realize how dangerous it can be, both for us and for the client himself, if we foster in any way the desire to "sell" focussed group research for purposes where it is not appropriate, or if we allow the insights drawn from focussed group research to be used as the kind of incontrovertible "proof" it is not intended to be.

My particular assignment for this conference pertains to another aspect of why clients are becoming increasingly fearful about the use of focussed group research--the quality of the focussed groups they have been buying and the absence of any standards or yard-sticks with which they can evaluate what they have bought.

In this area of research, in particular, a client is all too often buying the proverbial "pig in a poke." He has been told that qualitative research is "subjective;" that its value depends on the interpretive skill of the practitioner; that the practitioner has the special sensitivities and the special kind of knowledge that equip him to create the "magic" of the group interview and to draw marvelous wisdom from it.

But how, the client can argue, can he develop the confidence he is supposed to have; how can these special skills be measured or evaluated; what "checks" or controls are there; what ground rules ought to be met?

Today, unfortunately, there are no ground rules. In fact, within the past few years, I have seen hundreds of new practitioners of the "art" of focussed group interviewing spring full-blown on the scene with no more experience than that they have watched a few groups and are convinced that they can, with no difficulty, do equally well. In the course of a recent talk I had with one of my clients, we went through a list three pages long of research suppliers who had done focussed group work for his company in the past year, and neither of us recognized more than five of the names as trained, experienced, proven practitioners.

A further, very serious complication is the fact that there have never been any prescribed procedures or techniques which qualitative researchers have been required to follow. There are, in fact, almost no textbooks or training materials with which to help a person coming into the field learn the basics of the focussed group interviewing technique. As a result, the sad reality is that every moderator walks into the focussed group room and sets his own rules. It has even come to pass that any practitioner who manages to "discover" a "new and dynamic" gimmick that will make his groups seem different or innovative becomes this year's "hot" researcher.

While I certainly cannot deny that there are always, and should continue to be, new techniques to be learned and new skills to be developed with which to improve the success and effectiveness of focussed group research, I would like to make a plea for some kind of basic framework, some kind of essential starting point that can give this amorphous "discipline" some much needed discipline.

Over and over again, I have attended and even participated in conferences or panels on focussed group research when virtually every practitioner in the room had his own point of view, his own rationale and his own conviction that whatever he was doing was inordinately effective and successful. Yet another practitioner was taking a completely opposite position and advocating a totally contradictory approach. I remember a client sitting next to me at one of these meetings who left the room saying, "If I ever considered using focussed groups, this would certainly convince me to forget it. What they're trying to tell me is that 'anything goes' and that can't be the way it is."

I, for one, don't believe that is the way it is, and certainly not the way it should be. That is why I would like to spell out for you some of the essential ground rules I try to bring to bear and some of the techniques which I have been applying to focussed group interviewing for a host of years and literally

thousands of interviews.

I recognize, as I have suggested, that other practitioners may not concur in all of these points of view; also that there will certainly be approaches or dimensions that they feel are important which I will not cover. However, I would hope that this discussion can possibly become a starting point from which to build some body of basic information on the practice of group interviewing as it is applied in advertising and marketing research.

The first essential for effective group research, in my mind, is:

...KEEP THE CONSUMER IN A CONSUMER ROLE

I feel that it is crucial that the respondents in a group discussion at all times stay in their role as consumers. I believe that their only value in advertising or product research is as consumers, and that the only insights they can give us that will be of value must come from their experiences and reactions as consumers.

One of the greatest failures of qualitative research, in my opinion, is that the consumers who are called in for qualitative research projects are so often placed in the role of advertising or marketing experts. They come to an impressive, big company setting, are placed at an official-looking conference table, presented with a group of alternatives, and asked to make choices, judgements, comparative decisions, which they are not qualified to make and which are rightfully the province and responsibility of the marketing executive.

The only role that consumers should be asked to play is their own, and everything possible should be done to foster that attitude and to keep them in that role.

Our own way of doing this starts with the first contact with the potential respondents. The recruiters who solicit respondents set the mood by explaining the importance of the person's experiences as a consumer and the value of sharing those experiences with the people who are bringing them the products they use and trying to give them the things they want.

This same kind of positioning is also employed at the session itself when we make our introductory remarks. The participants are again told about their value as consumers and reassured about the contribution each has to make in drawing upon his own experiences.

We also ensure that the setting will not encourage an inappropriate role. Our sessions are held in a modest, "lived-in" looking living room located away from the rest of the business of the agency. The respondents sit on couches and chairs, never at a conference table. There are no obvious mirrors, no cameras, no video equipment. We studiously avoid these kinds of equipment because we feel they are overt and disturbing reminders to the respondents that they are expected to be "on."

We want to see our consumers and talk to them in a setting and mood that is as normal and familiar for them as we can make it because,

again, we want to be able to capture the experience and feelings they have in their every-day environment. We also want to establish the kind of rapport with them that will further foster this kind of relaxed and natural response.

This, then, brings us to my second essential for effective group interviewing:

...PROVIDE A MODERATOR THE CONSUMER WILL WANT TO TALK TO

To do this, we select and screen our moderators with scrupulous care, emphasizing always the warmth, the genuineness and the sensitivity which the moderator will bring to bear.

Although it is highly important that the moderator in a group session be non-directive and that he studiously avoid bringing his own influence on the responses, the personality and the demeanor of the moderator can unquestionably be a key factor in the effectiveness and even the "validity" of what comes out of the session. (I use the term "validity" here to reflect the honesty and openness of the respondents, and not in any statistical or quantitative context.)

Our experience has shown us time and time again how much more involvement and participation occurs when the moderator is warm and reassuring, when he (or she) has a quality of genuine interest and empathy with the respondents, when he is truly a listener, when he can readily communicate the sincerity and concern that make the respondents want to talk.

We also look for moderators who are "alive," who can keep the respondents alert and involved; moderators who can throw themselves into what is happening and can encourage the respondents, by their own sense of involvement and interest, to recognize that something interesting and exciting is happening.

A focussed group discussion should be exciting for the moderator, for the respondents, and for the listeners, and often it is the moderator who is responsible when it is not. But, even more important, if the moderator has not kept up the enthusiasm and interest of the respondents, the likelihood is that he will not be able to draw out the full range of insights the respondents can offer.

"Why bother?," the respondents might well feel in such a situation--and clearly, the research must suffer.

Essential number three:

...SET A TONE THAT MAKES THE CONSUMER WANT TO TALK

The success and effectiveness of the session is frequently directly related to the "tone" established for the discussion. The philosophy of interviewing to which I subscribe calls, as I have indicated, for a relaxed, informal mood and a free-flowing conversational feeling to the discussion. I believe that the goal of the session should be to draw out those responses and reactions and those insights which would appear if the participants were discussing the subject with their own

friends or neighbors, rather than any kind of formal, structured, question and answer situation.

We are, therefore, particularly concerned about the way in which questions are asked or probing is undertaken. The very same probe, we find, can be posed with so different an attitude that the response is almost a total reversal of the answer that might have been drawn if the attitude and the approach of the interviewer had been different.

We feel that it is extremely important to be supportive and reassuring in drawing out respondents. We feel that we can probe almost endlessly and still not "lose" our respondents as long as they recognize that we are not attacking or challenging or trying to make them look ridiculous.

Even when we are pursuing a point fairly relentlessly, we try to do it in a natural, empathetic way and our questioning becomes a kind of interplay of thought rather than an attack. We will even, at times, reassure our respondents with a comment like, "I think I understand, but could you explain a little more?" We also utilize the values of the group situation by calling on others in the group to lend their support by asking, "How about someone else? Perhaps you have had that kind of feeling (or experience), too. Can you tell me a little more about this?"

Sometimes we will be overtly apologetic, explaining that we don't mean to hammer at someone, but we truly want to understand.

We firmly believe that we can get more from respondents who want to talk to us and who are encouraged by our genuine interest to search more deeply into their feelings and emotions.

There are other practitioners who claim that they only get to the "real" stuff when they make the respondents angry, and they studiously go about challenging, provoking and even humiliating the group members. I have never yet seen a consumer buy a product in anger, nor have I seen him respond at the cash register because he is lashing back. And, in my book, the name of the game will always be to try to capture in the group session the emotional frame of reference the consumer is going to experience in the actual buying situation.

Our approach to interviewing also avoids an overtly authoritarian stance, and we attempt to achieve as much spontaneity and free-flowing conversation as possible. At the same time, however, we recognize the very important need for the moderator to stay in enough control that the discussion does not wander off in a totally meaningless direction. To achieve this kind of delicate balance requires essential number four:

...THE MODERATOR MUST MAINTAIN HIS ROLE, TOO

The skilled moderator knows exactly when and how to foster a free-flowing discussion while yet interjecting a new question or a new probe at just the propitious moment, both to maintain the role of moderator and to turn the conversation back to its appropriate course.

While I strongly prefer a non-directive approach to one that is restrictively authoritarian, I have seen non-directive interviewers who might just as well not have been in the room at all. And dreary, wasteful sessions they were indeed!

A middle course seems to work best for us. The respondents are encouraged to ventilate all their thoughts and ideas in a spontaneous manner and group interplay is encouraged at all times. But, the moderator is always on top of what is going on and the group is always aware that he is there to keep a structure.

Good group moderators, in my opinion, are like good teachers. They must earn the cooperation and the respect of the group members, and, having gained that good feeling, will have no trouble in maintaining both their own role and the proper function of the session.

The good moderator has also done a great deal of homework. He has had comprehensive discussions with all the people concerned with the project so that he understands totally the objectives of the research; what hypotheses have been raised; what the special areas of concern are. He has also given a great deal of thought to how to approach these various matters in the discussion.

Essential number five, therefore, is clearly,

...THE MODERATOR MUST HAVE DONE CAREFUL HOMEWORK

I can't stress strongly enough, that conducting a focussed group interview involves a lot more than just going in and asking a few questions. In fact, anybody who goes into a focussed group with such an attitude is running the very serious risk of being grossly misled.

The kind of question that is asked, how it is worded, when it is asked, how it appears in the context of the prior discussion, and, as mentioned before, the tone in which it is asked can all have a crucial effect on the kind of response it will bring. Serious, dedicated moderators will consider very carefully the interviewing guide they will use in a session and in introducing areas for discussion. They will work out in their minds the wording that will be the most non-committal and allow for the most meaningful kind of response from the group members.

We, for instance, never go into a session--even though we do many hundreds of sessions a year--without first preparing a written guide which we have thought and re-thought in terms of all its possible ramifications. We will carefully avoid asking a question like, "We would like to have your impression of this commercial," which, you can believe, gives you just exactly that: the consumer's own brand of strictly intellectual advertising expertise.

We've seen it happen over and over again: ask a question that makes a respondent an advertising expert and you've lost his value as a consumer.

We will also debate long and hard about what we are going to introduce first in our

questioning and what effect that might have on the areas that will follow.

Often, for instance, a client will suggest that we have what he calls a "general discussion" before we introduce his concept or his advertising. In many instances, this would be an incorrect and misleading approach because it would create an unnatural positioning for the concept. When consumers see an ad on television or in a magazine, they have not been thinking about the category for an hour or three-quarters of an hour before. They haven't intellectualized all of their feelings and taken strong public positions which they might have to reverse after seeing the concept.

In many situations, therefore, we present a concept or an ad as the original stimulus for the discussion and are, in that way, able to get the consumer's fresh and spontaneous responses. Then, later, we are able to work from that reaction into a more general discussion of the category and how the product fits into the category.

In other circumstances, however, the goal of the research is somewhat different, and the key requirement is to get a basic orientation into the category without letting the respondents be aware that we have a particular interest or axe to grind. In that instance, we would work downward from the general to the specific. That approach requires a completely different line of questioning.

Another case in point: we often struggle for a long time to avoid questions that will be so direct, so suggestive that it is inevitable that we will be misled. When I hear moderators asking questions like, "Is there anything unbelievable about that commercial?" or, "Is there anything you don't like about the commercial," I am reminded of the caution I heard a long time ago when I was a young mother. If you say to the kids, "Now don't forget, I don't want you to put beans in your ears," you can be sure they are going to put beans in their ears even if they never thought about it before.

One thing I can guarantee--and I have countless incidents to back it up--if there is some thing genuinely negative on the respondents' minds, you'll hear about it all too soon, whether you ask or not.

And what about that question that every client thinks he wants to hear: "Would you be interested in buying the product?" Again, what a shockingly easy way to be misled! How much more meaningful it can be to let the consumers talk about the product in their own terms, to let them deal, in an open-ended way, with how it fits into their lives and their needs in the category; to let them demonstrate, through their spontaneous enthusiasm or lack of enthusiasm and the degree of emotional involvement which they manifest, whether or not the product or the concept is viable.

This is why you are doing qualitative research, and qualitative research means that you are examining the quality of the response rather than the response itself.

Although I have spent several of the previous pages talking about the need for a carefully developed and well researched interviewing guide, I am now going to also suggest that, once the moderator goes into the room, he must never fall into the trap of getting caught up in his guide and so dependent upon it that it stymies the free flow of the conversation. He should also carefully avoid referring to the guide while he is ostensibly listening to the respondents.

There is nothing more important in a group than to listen intently and fully, for it is, as I mentioned before, the interviewer's interest and involvement that encourage the respondents to go on. As in the spy stories, the moderator must prepare the guide with scrupulous care, study it with equally scrupulous care, and then, so to speak, he must eat it. By the time he goes into the session, the guide should have become so incorporated into his thinking that it literally has become a part of him.

This is the only way the interviewer can carry on the free-flowing, comfortable, natural "conversation" that makes the discussion valuable. He must be prepared to come into the conversation and open up areas for discussion at appropriate moments while yet not interfering with the spontaneity of the responses and the valuable exchange of dialogue between respondents which is the rationale for the group experience.

This, in turn, points up another essential which the moderator must bring to the session. He must be skilled in the handling of the group dynamics which will occur and know how to use them in a positive way. He must also know how to short-circuit possible negatives of the group situation and minimize possible contamination.

Essential number six, therefore, is,

...THE MODERATOR MUST BE SKILLED IN THE TECHNIQUES OF GROUP DYNAMICS

We, in the qualitative research field, are all too familiar with the criticism that is repeatedly leveled at focussed group research, and particularly with the most prevalent of all, the belief that group members merely pick up from one another and reflect the same ideas.

Several of the clients I have spoken with even take the position that, after the first one or two respondents have spoken, none of the other respondents can be considered "fresh." Some-one else I know, who is himself a moderator, thinks it is very humorous to remark that six focussed groups are really just six individual interviews.

While I know that a snowballing effect can certainly occur in the group discussion environment, I also know that it is possible to counteract and minimize such influences with the right kind of moderating and a sensitive repositioning statement by the moderator. Let me give you some examples. We find, for instance, that, as soon as we see an idea starting to "catch on," we can get back on course by asking, "What about someone else? Does someone else feel differently?" With this kind of reassurance, a person who has a

different point of view feels free to disagree and express his own position.

We also say, at times, in order to turn a discussion away from a course, "That's one point of view," or, "That's an interesting point; now what about some other thoughts?" Or, we call upon a respondent who has shown signs of another kind of thinking and bring him into the conversation. That then provides the moderator with the opportunity to say, "There seem to be several different points of view that have been brought up. Where are the rest of you on these? What has your experience been?"

In many instances, we also try to avoid group influence by going around quickly at the start of the session to get an immediate, first-thought reaction from everyone. We don't probe at this time or let anyone develop a point of view. Then we go back and pursue the various reactions more totally. But we have first had an opportunity to get all the first-blush ideas out on the floor.

These are some of the ways we use to avoid group influence when it is overt and easily spotted. There are, however, the problems that occur when the effect of a remark or an idea has sparked a whole line of reasoning that might not otherwise have occurred. Even the respondents themselves may not know that this kind of influence has been at work. To guard against this kind of contamination, we always insist on supporting groups, at least one and hopefully more, with essentially the same types of respondents, so that we can find whether the same ideas and reactions repeat, even though the conversation may take a somewhat different turn.

Another much talked about and highly exaggerated hazard of the group interviewing situation is the dominant respondent and/or the non-stop talker. Again, this can only occur if the moderator permits it; and a skilled, competent moderator knows how to deal with such offenders. They require firm and direct intervention, albeit intervention which is civil and good-natured and which doesn't embarrass either the offender or the other members of the group.

If a disruptive or dominating respondent is put down too sharply, other members of the group might be fearful of exposing themselves to the same kind of attack, so they take a safe course and say nothing.

The best technique is usually to break right into a sentence with a reassuring, "That's interesting, but let me just find out what these people over here have to say," or, "Yes, I understand. Let me get some other opinions."

The moderator must also studiously avoid involving the big talker any more than necessary in other parts of the conversation and should direct questions or comments away from this person and toward others in the group wherever possible.

The eager talkers are more often than not counterbalanced by reluctant talkers who need special encouragement and a show of interest on the part of the moderator. It is interest-

ing to see how easily they can be brought into the conversation with reassuring support and some show of recognition from the moderator like, "What about you? You've been very quiet, but I'd really like to know how you feel about some of these things we've been talking about. What ideas did you get when you saw the commercial?"

- The moderator who is skilled in working with group dynamics knows how to draw out individual experiences and reactions.
- He is adept at "using" what is happening in the group to stimulate and broaden the discussion.
- He encourages positive interaction between respondents and recognizes the potential for them to spark one another into deeper and more fruitful introspection. At the same time, he is ever watchful of negative interaction and is quick to head off painful incidents that can interfere with the freedom with which all the respondents will react.
- He knows how to recognize the difference between the intellectual and the emotional-- between what the respondents are saying and what they are experiencing.
- He is constantly aware of the mood and the demeanor of the participants and knows that the strongest "clues" in "reading" the group will come from what is happening and from the emotional affect he perceives in the respondents. He is always listening with a particularly keen "third" ear.
- He knows how to cut right through the expertise and get the respondents back on the track.
- And, most important of all, he knows how to interpret the group for his clients so that they, too, will be able to take what is valuable and genuine from the group.

Basically, the good moderator is a "pro," and he got that way by learning his trade well and long.

Again, I come back to the very same point at which I began. We need more practitioners who are indeed "pros" and who have learned their trade well and long. We need schools and work-study programs where they can be trained and developed. And, we need, desperately, some kind of quality control.

If we continue in the haphazard, everyone-set-his-own-rules manner we have been following, the inevitable result will be that more and more clients will become disillusioned and even embittered about this valuable, and in many cases, irreplaceable research tool.

Some clients already feel that way. Others are moving in that direction. And what a crying shame that is!



PRESEARCH AS GIRAFFE\*: AN IDENTITY CRISIS

Jane Templeton, Ted Bates, Inc.

For the past ten months -- give or take a month -- I have had on my desk a personnel form requesting me to produce a job description. I've gotten as far as name and title, but the twelve lines (I've counted them) provided for the actual descriptive meat remain blank. Given the usual personal vanity about compressing one's vital and complex function to a simple objective account, and allowing for an innate tendency to procrastinate, ten months is still a long time. Once, out of impatience with the delay, I decided that a flat statement of the operations that fill the time might satisfy, and started with "I conduct and interpret group interviews". That leaves out meetings and white papers, but does cover the bulk of time and motion. However, it dawned on me before the crossing of the first "t" that it was a cop-out and I reached for the O-Pake. To say that I conduct and interpret group interviews is like a tailor saying that he threads needles.

It would be equating, e.g., a panel recruited for casting purposes with a fishing trip for motivational dispositions in a new product area. In casting interviews, we want to know how the panelists feel about particular product-connected areas, and to elicit this information, if possible, without ever touching on those exact questions (so as to preserve some degree of freshness and spontaneity when the studio interviewer asks them). We must also be able to gauge whether, for any particular panelist, the vivacious and responsive lady or gentleman we see in the group interview will continue to scintillate under production stresses, or will go blank or become afflicted with "peanut-butter-mouth" in the studio situation.

When we are trawling for ideas and needs associated with a new product, we are similarly wary of introducing key topics -- but not because we don't want them discussed. We try to encourage the respondents to raise the relevant issues, so we can discover what they are and where they most naturally "fit" into what these respondents currently think about. We try out various broad general areas of discourse with some likely relationship to our product -- working from general to increasing particularity from the responsibilities of the homemaker, say, to nutritional accountability, to breakfast, to breakfast juices .... until somebody's switch is tripped. But once this happens, we pursue

---

\*My first job in advertising was an act of faith on the part of the research director. He coined a title for my function and assigned me to the creative group. When I asked what I would do, he said: "You've done group therapy, haven't you? You should be good at focus groups. They'll tell you what they want to know, and you'll interview a group of consumers and come back with the answers." Aside from instructions on recruiting protocol, this was the sum of my instruction. Soon after, at an intra-agency presentation, I asked one of the art directors to do a visual flip-chart for me: two giraffes, behind a zoo-enclosure; one of the giraffes has bent his neck over to read the sign, and is saying: "We're Giraffes." Many years later, having watched and listened to many other researcher's conception of "focus groups", I find I've come full circle and am back to the Giraffes.

the exploration of opinions, attitudes, and feelings as deeply as the group will go (and group tolerances differ) to find the points at which fundamental internal states intersect with behavior.

And what about creative reconnaissance, which is another challenge altogether? Where to put new business sallies which -- with half an eye to the theatrical use of the videotape -- call for on-the-spot interpretations? Manifestly, naming the tool isn't going to get me off the hook. I could as usefully say that "I do marketing research".

What is it in general, then, that distinguishes the functions of Presearch from the inbaskets of the account research people down the hall, with whom we frequently share projects?

- o There are the apparent, easily observable distinctions that Presearch is not (as the account-research people are) affixed to any particular account group, but float from account to account as need dictates. And our "data-gathering" differs in that most of it is done in the agency's "fishbowl", or in similar one-way mirror facilities outside the city. But these distinctions are really extraneous.
- o Paradoxically, the essential difference between what "we" do and what "they" do occurs during the one stage in the project when we all appear just the same way: shuffling papers, frowning, and getting cranky about interruptions. We are all "interpreting" data, but Presearch deals, in a different kind of interpretive responsibility than that assumed by our partners at the other end of the floor.

For example, if someone is not copied on a survey report, or left his copy on the New Haven train, or takes issue with the conclusions it propounds, he can always go back to the raw data, to seek his (her) own conclusions. Our "data" are only data to us, and our clients, in or out of the agency, pretty well know it. They know it because:

- o We've worked hard to help our internal "clients" in the creative and account divisions to recognize the complexity of group interview data, and many of the agency clients now understand this as well.
- o Also because some of them have learned by experience that our interview materials do not lend themselves to literal interpretation. A typed transcript of a session, for instance, often has the same relationship to the interview guide we originally set down, that a 36-in snake has to a yardstick -- it's hard to line up and check off. The topics suggested in the guide are raised, but not sequentially, and topics tend to be interwoven and reappear, sometimes with very different implications, throughout the interview. Further, the written material may contradict the transcript, or both may be belied by the behavioral observations or by the projective data.

The focus of this paper is interpretation of group material, rather than details of procedure. But at this point, a general summary of our particular m.o. -- what

we do do and what we don't -- may prevent confusion.  
Presearch group interviews:

- o Employ a relatively formal setting: a conference table in a large office within the agency (or similar facilities in other cities).
- o Employ a one-way mirror and/or videotape, and overhead microphones which are connected to an audio recording system and to a loudspeaker in the observation room behind the mirror.
- o Are conducted with panels of, typically, 10 - 12 consumers, using the following format:
  - After panelists enter and seat themselves, and are given refreshments, there is a short Warm-Up, during which everyone including the moderators introduces him(her)self to the rest of the group and "ground rules" for the interview are stated.
  - This is followed by a Predisposition discussion, which concerns itself with the contexts in which the product (we are to explore) is bought, used, and thought about. This will include general reactions to advertising in the product area generally.
  - We then introduce materials: concepts, rough or finished creative executions, products, etc., and ask panelists first to write, privately, their immediate reactions to each of the materials, and then to discuss it. This pattern of "write, then talk" is continued until all materials have been exposed.
  - After all materials have been discussed individually, there is usually a collective and comparative discussion of everything exposed to the respondents.
  - The discussion ends with the Wrap-up: a summary statement of what panelists think the group as a whole has expressed during the interview.
  - Before leaving, panelists complete a brief demographic questionnaire and a self-administered projective instrument (drawings and stories).

In view of the several papers delineating rules and standards for moderating group interviews, to elaborate further on our particular interview protocol would be redundant and, worse, presumptuous. What I am describing is in no way intended as a prescription for "how to do group interviews". It is only a statement of how we do group interviews.

But a clear understanding of the nature of our interpretive grist does require more detailed description of how we conduct our sessions. One of the essentials of our manner of approach is the avoidance of direct questions in all parts of the interview. This type of interchange happens as near to never as conversation permits, which accounts in part for the non-sequential, interweaving flow of discussion. Obviously a direct question procedure would be simpler both to moderate and to interpret. But we feel that answers to direct questions are dangerous. They cannot give us some of the kinds of information (motivational, qualitative) that we are seeking, but more importantly, they tend to provide answers which can seriously mislead us. We have defined four reasons for avoiding the hard-edged frontal question:

- o Partly, we eschew direct questioning because this kind of interaction is boring. It produces emotional disengagement from the topic, for everybody concerned, to the point of automatic, unsearching answers. Feelings would still be going on, because feelings do operate constantly in people. Some of the feelings might even be strong (like the itch to get away or to gag the moderator), but the feelings each of the panelists might be experiencing could have low, intermittent, or no relationship to the stated topic in a direct question-and-answer interview. A clever and funny moderator can, of course, make even this format entertaining, but who's been interviewed?
- o Also, parallel questioning of individuals in the group is a very efficient anti-personnel weapon, in the sense of group dynamics. Respondents get disengaged -- not only from the topic, but also -- from the other people in the group, so that interpersonal provocation, influence, and drift are no longer discernible.
- o Too, asking a question directly does not allow the issues to emerge spontaneously. This deprives the moderator and the behind-the-mirror viewers of the opportunity to see its relative saliency, or to weigh and consider the company which that issue keeps: the ideas immediately associated with it, the feelings that accompany it, the language used to express it, etc.
- o Finally, there is a less obvious reason for avoiding direct questioning; less obvious, but central to the different kind of interpretive responsibility assumed by the group-interview researcher: we avoid direct questions because of the difficulties in willingness to respond. Not that the subjects are unwilling to answer direct questions. Rather, they are quite willing (providing the questioning hasn't put them to sleep). In fact, they are willing regardless of whether or not they know the answers. I'm not suggesting malicious uncooperativeness on the panelists' parts, nor am I falling into the trap of "insulting the intelligence" of our respondents.

Remember that the kinds of questions typically addressed by this type of research are uncommonly complex. We go into our groups committed to come as close as possible to answering a brain-buster like:

"If this storyboard is produced as a television commercial, what reactions would the people in this panel be likely to have to it, and what would they do about it?"

Questions like these are not only complex, but also require attitude projection and behavioral conjecture that I'm not sure anyone can manage accurately by introspection. Simple past, present, and future tenses are hard enough to introspect about, heaven knows, but our questions -- if directly asked of group panels -- would have to be set in some obscure tense like the pluperfect conditional.

"If such and such were to happen, in the following situational context, then would you...?"

Questions of this sort are asked, by the people who request the project. And they are duly set down in the "Background and Purpose" section of the final report. They are also, more often than not, answered -- by the person responsible for the project -- with limitations and caveats reflecting the size and probable biases of the sample. But they are neither asked of nor answered by the panelists themselves, directly. Because if we

asked, they would answer. And they not only don't know the correct answers, they don't know that they don't know.

That statement, and the claim which is implicit in it, bear some thinking about. Assuming that panel members want to be cooperative (and we usually assume that), and that they have had twenty or more years to get a good fix on themselves, it is a lot to say that we can learn things about them in a two-hour group interview that even the brightest respondent, with the best intentions imaginable, can't tell us. But we can -- and do.

For starters, the things we principally want to know about them are things that they rarely think about concentratedly: buying and brand behavior, usage, product attitudes, etc. -- things which are negligibly important to them as ordinary citizens in the real world ... and are the very essence of our real world. They have little motivation to search themselves for better understanding of this sphere of their lives. Our motivational stakes in understanding these things, on the other hand, are very high indeed. So we'll try harder.

Also, we bring to the interview situation two kinds of expertise which the panelists don't have. We use our expertise in human behavior and marketing strategies to figure out how internal events like feeling, attention, and memory combine in the ultimate sacrament of reaching a hand into a pocket to buy our product. So when we talk about "interpreting" consumer reactions to get the answers to specific questions or to make specific recommendations, we mean something different from the "interpretation" of a questionnaire survey.

In interpreting survey data, the respondents' actual statements are treated as factual, and interpretations are based on measurements and comparisons of these "facts". In the case of group interview interpretation, "what they say" may be amended or modified, or in some extreme cases, even totally contradicted by the interpretation.

I don't mean that the panel's reactions are ignored. On the contrary, all of the respondents' communications are taken into account, both in aggregate and minute-to-minute. As we perceive it, at least three communicative channels are open either all of the time or intermittently through the interview, providing us with three kinds -- or levels -- of information:

- o The level of public affirmation: This is what panelists actually say. It is their interpretation of what they think and feel, impacting with the social role they are trying to maintain, in conjunction with the expressed views of other panel members. We haven't asked them to go "on record" with flat "yes"es and "no"es, so they are not greatly concerned with consistency (or can't keep track of it). Neither are we. We watch motivational drift closely because an about-face is as useful to us as a to-the-death stance. The language they use in discussing what we are there to talk about is also a part of public affirmation, and since they have usually introduced the topic, their language is relatively uncontaminated by our expectations of how they will talk about it.
- o The level of private acknowledgement: This is what they write, on our open-end questionnaire, as soon as they have been shown a commercial, or storyboard, or concept, and before any discussion takes place. If your eyebrows go up about the assumption of some independence between these two levels, I have no hard-headed answer. We do ask everybody to turn the written forms face down before open talk com-

mences, and we begin discussion in a different way than the questionnaire, so that there is no exact parallel. But we can't erase the written answers from their minds. Exactly how or why it works, I'm not sure, but that it works, I'm pretty confident about. Written reactions very often sound as if somebody else came in to write them, when compared to the group interchange, and they rarely track very closely the direction of open discussion. We think that the written material reflects what they think "in solitary", as opposed to feelings they subsequently "own" under social pressure. Or perhaps the spoken comments mirror more what they want to be heard saying. Fitting written statements into our interpretive scheme, we use content plus indications of intensity of emotion or opinion (underlining, exclamation marks, heavy pressure) and involvement (how much is written, signs of personal projection of product use, etc.)

- o The level of personal revelation: This is first of all, what we learn from their non-verbal communications: vocal range and variation, postural changes, facial expressions, constrictive or expansive demeanor\*. Three respondents can say the same thing and express quite different inner states. Consider the phrase: "Frankly, it leaves me cold." Assume that one respondent making this comment spits it between clenched teeth, leaning forward, hands gripping the edge of the table; that a second panel member says it in a low voice without inflective melody, suppressing a yawn, leaning back with her hands slack in her lap; and that a third panelist says it almost laughingly, sitting forward hugging her arms, maybe with her hand across her mouth, swinging her chair, and with her eyes sparkling. It's up to the moderator to be aware of such behavioral distinctions, not only in the person speaking at any one time, but in the group as a whole. About a group, we may note its speed of warm-up, whether -- and when -- they are autonomous or look to each other or the moderator for guidance, the intensity of controversy, the tendency to return to -- or to avoid -- particular topics.

The figure drawings and stories we regularly ask panels to produce are also sources of the "personal revelation" level of data, and are used to elaborate, underline, or reconcile sketchy, ambiguous, or contradictory impressions.

It will have occurred to you, of course, that when we shift our attention from one level of response to another, during the interview, we are confounding the distinction between moderating and interpreting. And of course, you are right. The line is a little fuzzy. Some clearly interpretive operations may go on -- even out loud -- while the interview is in progress. This is a decision based on moderator-judgement. Interpreting is almost sure to raise the feeling intensity of the group, and may generate a certain amount of anxiety. The balance is difficult to describe. Placidity, comfort, and consensus are by no means the only -- or even necessarily the most ideal -- forms or outcomes of a group-exploration. But on the other hand, the moderator assumes an obligation to keep feeling-intensity and interpersonal contention within tolerable bounds, and to hold rein on behavioral chaos. If the group is judged to be firmly-knit

---

\*If you were worried about the independence of written/verbal responses, you may be beginning to question what we use as a behavioral baseline against which interview behavior stands out and can be interpreted. We establish this informally, for the group as a whole and for individuals, during the warm-up.

enough to contain stress, and the individual(s) in questions is (are) self-searching enough not to become over-anxious, then it is interpersonally defensible -- and can also be extremely productive -- to offer interpretive probes like:

"You say that it leaves you cold, but I'm getting a very different message from your voice and manner -- that you really dislike it very much" (or "that something about it amuses or delights you"). Can you clarify those different communications for me?"

When the respondent can assimilate this degree of conflict and has the self-awareness to resolve it, or alternately, when the climate of the group is supportive enough so that other respondents will rush in to help one of its members to better self-understanding, these interpretive interchanges are not only helpful to the moderator (and viewers of the session), but also provide the respondents with the uniquely heady achievement of insight. This explains the apparent paradox that some interviewers which are apparently charged with ambivalent or unpleasant feelings and interpersonal strife are frequently received by respondents as joyous, uplifting experiences.

In the closing minutes, during the interview wrap-up it is customary for us to invite interpretation from the panel. At this point, the moderator says something like: "You've all been sitting at the table just as I have. If you had the assignment of summarizing what this whole group felt about \_\_\_\_\_, what do you think you'd say?" Because our sessions tend to be up-tempo, interpersonally active, even sometimes tense and factional, respondents usually want a degree of closure, and will generally jump in -- often spotting things that neither moderator nor viewers have picked up.

The first formal, unadulterated "interpretive" act takes place during our "post-mortem": a debriefing session which is held immediately after the interview, and is attended by the moderators, any viewers who have hung on staunchly to the very end, and others who are concerned with the project and may or may not have been able to watch from behind the mirror. During this informal rehashing, we note the conspicuous themes of the discussion, mention spontaneous impressions of the behavioral flow, perhaps glance at the written responses and the drawings, and negotiate a general sense of the direction of the group, using the contributions of everyone who participates.

Now the major job of interpretation begins. This is part of the job description that gives me the most trouble. It is also the part of the job that gives me the most trouble ... and the most personal "juice". Interpreting one or a series of group interviews places great demands on intuitive and organizational skill, and I never finish a report without the feeling that some small truth has been extracted from tons of pitchblende.

Obviously, the interpretation which is finally made will depend on the purpose for which the groups were scheduled and on the form in which results are to be communicated. But whatever the purpose and intended format of presentation, the act of interpreting our group interview data consists in the bringing together of disparate material (private, written reactions, interactive discussion, observed behavior, drawings and stories), weighing and sifting of all inputs, and organization of these multiple clues into an articulated set of premises and speculations.

To take one example, probably our most frequent assignments are aimed at assessing panelists' reactions to creative material. The way we collect data on the rel-

ative power or "goodness" of concepts or executions is very different from standard copy-testing procedures. We don't, for example, measure increments of interest and importance, nor do we construct scales. It follows that the kinds of answers we give to the questions asked of the project will be different kinds of answers, in meaning if not in labelling. For instance:

- o Comprehension: We do include in reports some estimate of how well our panels seem to understand the message in a concept or execution, but with one important difference. We assume that in any communication the message that is received is as valid as the one transmitted. If most of the people we talk to "understand" copy to mean the same thing, and if they are positively affected by what they think it means, we say that this is good comprehension. We will say this, even if what panelists understood is not what the copy meant to say (of course noting in the report that there is a gap between what we think we said and what they think they picked up).
- o Persuasiveness: We also milk the data for anything they can tell us about the extent to which panel members are convinced by our creative material that they should try the product. Precisely because we don't ask: "Would you buy?", we feel free to place some weight on spontaneous statements of buying intention, especially if they are supported by indications that the respondent has projected the buying or using of the product into his future expectations, e.g. by incorporating it into a larger plan: "I would buy it and use the money it would save me to go to the movies." We also watch -- and use -- things like switching from the conditional to the declarative mode: "I would buy it so that I will be able to save money and go to the movies with it." Facial expression and behavioral responses bear on the state of persuasion as well. Also fanning contention and controversy allows us to observe how much respondents who are persuaded will argue with those who are not, or how stoutly they will resist the arguments of those definitely opposed.
- o "Importance" and "Recall": We combine these ideas in a concept we have privately labelled "Embeddedness". This has to do with the extent to which other, subsequent life events which these particular respondents are likely to encounter will tend to evoke rather than to bury their recollection of the message (the "message" being what the respondent got out of it). If, e.g., the next time the respondent gets into a lather about rising prices, she'll probably remember our product, and if she's a type who padlocks her purse, and lathers often and intensely, then for this respondent, the message is highly "embedded".
- "Embeddedness" also includes the quality of "identification" which panelists may feel with the idea, the situation, or the people depicted in the execution. Someone who sincerely gets a shock of recognition: "Hey, that's me" when he looks at a commercial is apt to be reminded of the execution everytime the parallel situation occurs in his own life.
- o Believability: Is something that shifts significance according to the product, the degree of belief or unbelief, and the reason for which it is believed or not. We report on it, when it seems to be important, but it has no permanently assigned evaluation. Clearly, a cosmetic product that is "too good to be true" may have created a very positive impression, while "some of your damned advertising doubletalk" in connection with bank services or a food's freedom from additives has not gone down too well.

- o Liking: The romance which panelists have with the creative material is by far the untidiest of the creative responses we consider. Presumably, an ad or commercial that isn't sufficiently "liked" won't be allowed to deliver its message. On the other hand, we've all seen "adorable" campaigns that didn't move the product, and "outrageous" ones that did. We have to address "liking" in reports, because it is something that panels talk about, but there is no one standard rule for interpreting what "liking" a commercial or advertisement has to do with purchasing the product. We tend to think that strong feelings in either direction register more clearly and last longer than the most benign low level response. A comment like: "It's short and to the point and no-nonsense" -- whatever pejorative connotations are intended -- is nearly always a kiss of death.

(though possibly less simple) inventory of one's own skills, blind spots, biases, and expectations as possible. This will instruct the professional as to what he can do very well, where he must exercise caution, and -- when the data are before him -- help him to recognize a moderator-skew when he sees one.

As for how we go about combining our three layers of interview material into estimates of persuasiveness or embeddedness or whatever other judgements we have been asked to make, there are, again, no invariable rules. When written reactions are at odds with socially aired opinions or feelings, we can't assume a priori that one or the other is the more "true". We must take into account the type of product, the experience of using it, the probable impact of social pressure on the product category, and so forth. If the written comments are more positive than the beginning of the discussion, it may mean that respondents are drawn to the product, but must pay lip-service to consumer cynicism. If attitudes expressed in the group tend to become more positive as the discussion continues, we would probably assume that this was true. On the other hand, initial private acceptance followed by public rejection may equally well show a quick disenchantment with advertising claims perceived as superficial or irrelevant, and in such a case, could indicate more intense net-aversion than when both written and verbal responses are moderately, uniformly negative.

There are no formulae. There are, alas, few precedents. Sometimes, long familiarity with a product or product group will give us a reassuring feeling of solidity, and some readymade hypotheses for explaining contradictions. But even here, we must be alert for signs of change. There is also a kind of cumulative serendipity that permits us to recognize in one product category attitudes that are familiar from another, and to speculate about whether, e.g., a product that used to be almost purely cosmetic is beginning to shift to a medicinal image (since attitudinal patterns are suddenly similar to those habitually seen in drug product interviews).

By-and-large, however, once the interviews are done and the various interview products sifted through, we are alone with the data and whatever tools we have acquired to organize them into a final report. Partly, I am hindered in delivering an adequate job description because I feel I should append a resume. It would be difficult to find a vocation that challenged more completely the sum of knowledge and skills I can muster. My group interviews -- and the reports that summarize them -- are as they are because of my academic and clinical background, and have gotten better as my marketing background increased. Group interviews, generally, have astonishing flexibility, and can absorb whatever one brings to them.

Having begun with the premise that I was setting forth one way -- not the way -- to use group interview research, I find I do have something to say about how one "ought" to approach group interviews. One should approach them with as simple and clear an idea of the objectives as possible, and with an equally clear

WHAT ADMINISTRATORS SHOULD KNOW  
ABOUT THE GROUP INTERVIEW

George J. Szybillo, New York University

Administrative responsibilities are often cited as including problem definition and specification of how information relevant to the problem can be acquired. Regardless of umbilical association (e.g. consumer marketing, industrial manufacturing, government, non-profit organization), administrators have similar needs with respect to these responsibilities. The group interview -- the direct verbal interviewing of respondents in groups rather than individually -- is a research technique with implications for each of these responsibilities. Group interviews highlight what factors are associated with problems and often suggest strategies for how information on these factors can be collected. Group interviews are descriptive, focusing on words rather than numbers. They describe the what and why of a problem with little concern for how many or how much. Emphasis is placed on the quality of responses rather than the quantity of responses. How does this principal characteristic of the group interview relate to the functions for which group interviews are used? What procedural steps in group interviewing should the administrator be aware of? What are the advantages and disadvantages of this research technique?

#### Functions

There are at least three general functions which the group interview may serve individually or collectively. Goldman (1975) suggests that the qualitative aspect of group interviews may address consulting, creative or research needs. His categorization suggests that the qualitative aspect of group interviews may address a multiplicity of needs. His categorization suggests that group interviews do not necessarily have to serve the purpose of further research.

Administrators are often confronted with problems with which they are unfamiliar. In these situations the administrator's principal objective may be to learn about the problem. For example, consider the product manager in a package goods firm with responsibility for rice products who is to be promoted to group product manager for staple products. However, the firm presently has no staple products other than rice. Our manager needs an understanding of what factors are perceived as assets and liabilities of present and, perhaps, potential staple products. Where a thorough orientation to a new area of inquiry is the administrator's principal objective, group interviews can be said to serve a consulting function (Goldman, 1975).

Group interviews can also serve a creative function. Holtzman (1975) illustrates that copywriters often need guidance on not only what to say in an advertisement, but how to say it. Early quantitative research for the women's shaver market stressed the importance of functional characteristics of electric shavers (e.g. handling, closeness of shave, etc.). Yet the advertisements stressing these features in conjunction with femininity and daintiness did not stimulate the market for years. Group interviews for Remington indicated that a large part of the women's shaver market wanted not only functional features, but some resemblance to the male counterpart. Emphasis was changed in the message context from femaleness to maleness and sales increased.

Group interviews can also serve to directly and indirectly guide quantitative research (Johnson, 1975; Dupont, 1975). In this function they can: 1) suggest or reconfirm hypotheses to test, 2) provide language for question structure which is most likely to be familiar to respondents, and 3) suggest advantages and disadvantages associated with other data collection methods. For example, Dupont (1975) found that approximately 1,000 recent new car buyers had to be contacted by phone to recruit 20 group respondents of Volvo considerers. Phone or personal interviews were therefore judged to be beyond the cost Volvo had anticipated for the quantitative phase of the study and direct mail was used.

#### Procedures

##### Preparation

Administrators should be involved in defining the problem to the group interviewer or moderator. This is important for the moderator not only interacts with respondents, but interacts with them for a purpose. Mutual understanding between moderator and administrator clarifies, why the study is to be done, what specific areas are to be covered, how the results are to be used, and how and to who the results will be presented (Payne, 1975). After defining what questions are to be addressed a guideline is frequently developed to place questions in a logical sequence and to provide a frame of reference for how both the administrator and moderator believe the group should proceed.

An administrator's involvement with the project should not stop at guideline preparation. He must specify who he wants in the group and who he does not want in the group. Considering the first point, a baby food manufacturer once used the description "mothers with young children" and found that the recruiter had assembled his group respondents from a list of Leche League members. These women breast feed their babies and were not respondents the manufacturer had hoped to address. This example illustrates that recruiters often use sources other than telephone or door to door recruiting. Such variability in recruiting practices leads to variability in recruiting costs. Lists of respondents also may lead to groups composed of "professional respondents" and "friends". Neither of these types of groups may be acceptable to the administrator. Further exclusions could include marketing research, advertising agency, and competitor or corporation personnel.

Group size and number of groups should be of concern to the administrator. Most facilities and moderators seem to accommodate groups of seven to ten respondents (Payne, 1975; Wells, 1975). It is a good idea to over-recruit for groups. At any time respondent concerns (e.g. little Johnny has a runny nose) can become more important to the respondent than an administrator's research. The number of group interviews varies depending on the scope of the study. Generally, analyses are performed on two to four group interviews. Studies which demand analysis for different groups determined by a number of factors (e.g. age, income, marital status, etc.) necessarily involve more groups.

## Moderation

There are different interviewing styles for the moderation of groups. Aspects of nondirective interviewing have been detailed elsewhere (Goldman, 1963). Both Axelrod (1975) and Wells (1975) have discussed aspects of the directive style. Elements in a directive oriented introduction may include: overview of where the discussion will go, instructions on participation, stress of importance of reactions as participants, and a request for general background information on participants. Each of these elements deserves further comment.

By presenting an overview of the discussion outline the moderator gives the participants an idea of the scope of the discussion and the boundaries associated with the discussion (Wells, 1975). Instructions requesting that respondents speak up highlights the tape recording of the session. This can be explained by indicating that the respondents' responses will remain anonymous, and that the tape is used only to facilitate later analysis. It also is easy to explain in this context that while spontaneous reactions are appreciated, multiple conversations within the group, at any one time, will mitigate the usefulness of the tape. Stress on participant reactions, as participants, informs the respondents that their own reactions are sought; not their interpretation of the problem as they would adopt the role of a manager or advertising executive. Request for general background information from the respondents enables them to speak first on a subject they should know well; themselves. It also allows them to warm up to the moderation process and gives them an idea of how the moderator may lead the discussion.

Axelrod (1975) notes that the moderator should set: 1) a tone that impels respondents into the discussion, and 2) a pace that allows covering the specific areas of research interest. Emphasis should be placed on probes which simulate an interplay of thought rather than an attack (Axelrod, 1975). Administrators may argue for the same conclusion; "silent" research helps no one.

Respondents are generally informed of the duration of the interview (e.g., 2 hours) prior to participating. Therefore, to allay rushed answers, the interviewer must pace the session so that areas of interest get appropriate "air time". In the most desirable type of interview, the moderator also balances "air time" for respondents. This means that the moderator should be capable of eliciting responses from reticent group members and thwarting dominance of discussion by the more aggressive group members. The first situation may be alleviated by such probes as "What about someone else?" "What do the rest of you think on that topic?" The second situation may be alleviated by the same probes but tactics such as breaking the respondent's train of thought or nonverbal presentations (e.g., moderator boredom, avoid eye contact) may be necessary.

## Analyses and Report Preparation

Analyses of group interviews and report preparation can range from a one paragraph account of what prevalent verbalizations occurred in the groups to a multi-page report describing and interpreting multiple levels of group responses including verbalizations. Where the administrator has had intensive participation in the study and when respondent reactions appear to be uniform, a report beyond the obvious may not be desirable. On the other hand, where administrator participation has been low, respondent reactions diversified, and organization of previous research findings poor, an extensive, integrated and interpretive report may be called for (Templeton, 1975).

Extensive reports generally include discussion of what was said, how it was said and description of respondent reactions when it was said. It becomes readily apparent at the analysis stage that if similar research topics were not discussed across groups, few comparative statements can be directly made. A moderator with good topic area control of "air time" is greatly appreciated here.

Dr. Jekyll and Mr. Hyde

As with most things we can discern good and bad characteristics. The following can be considered disadvantages and advantages associated with group interviewing. Among the disadvantages are: 1) the findings of group interviews are often not projectable, in a statistical sense, to the population from which they were recruited due to limited number of participants and non-random processes associated with recruitment; 2) it provides less detailed information on a person than the individual interview; and 3) group respondents may overly favor concepts or objects presented to them if they perceive them as being associated with a friendly moderator.

Among the advantages are: 1) group interviews can be conducted very quickly and at relatively low cost; 2) group interviews provide the opportunity for the research professional and administrator to directly observe respondent reactions to the problem; 3) group interviews are more flexible than many other research techniques; 4) group interviews expose the contingencies on behavior or the dynamics of decision-making; 5) group interviews often foster the interchange of stimulating ideas; and 6) group interviews provide information to an administrator in a form that they can understand and appreciate. Rather than a basket of percentages and unexplained tables, typical of many survey reports, the group interview report provides the administrator with the reality of how respondents view his problem in their own words.

## References

- Axelrod, Myril D., "The Dynamics of the Group Interview," in Beverlee Anderson (ed.), Advances in Consumer Research, 3(1976).
- Dupont, Thomas D., "The Exploratory Group Interview in Consumer Research: A Case Example", in Beverlee Anderson (ed.), Advances in Consumer Research, 3 (1976).
- Goldman, Alfred E., "The Group Depth Interview," Journal of Marketing, 26 (1963), 61-68.
- Goldman, Alfred E., "Development and Application of the Group Interview," in Beverlee Anderson (ed.) Advances in Consumer Research, 3(1976).
- Holtzman, Eleanor, "Love, Sex and Intent to Buy," in Beverlee Anderson (ed.), Advances in Consumer Research, 3(1976).
- Johnson, Deborah, "The Shopping Crisis: The Focus Group Method as a Way to Examine Changing Trends," in Beverlee Anderson (ed.), Advances in Consumer Research, 3(1976).
- Payne, Melanie S., "Preparing for Group Interviews," in Beverlee Anderson (ed.), Advances in Consumer Research, 3(1976).
- Templeton, Jane, "Presearch as Giraffe: An Identity Crisis," in Beverlee Anderson (ed.), Advances in Consumer Research, 3(1976).
- Wells, William D., "Group Interviewing," in Robert Ferber (ed.), Handbook of Marketing Research, McGraw-Hill Book Company, New York, (1974).

# PROFILING THE SENIOR CITIZEN MARKET<sup>1</sup>

Kenneth L. Bernhardt, Georgia State University  
Thomas C. Kinnear, University of Michigan

## Abstract

This paper examines the personal characteristics, shopping behavior, credit usage, media habits and leisure time activities of consumers in the senior citizen market. A comparison on these dimensions is made with consumers in other age groups with many similarities and differences noted.

## Introduction

The growth in the size of the senior citizen market has been well documented in the marketing literature (Salmi et.al., 1972; Goldstein, 1968; Forbes, 1969; Business Week, 1971). In 1900, there were just over 3 million people age 65 or older, approximately 4 percent of the population at that time. By 1950, the number of senior citizens had grown to 12.3 million, and the percent of the total population in over 65 age group had become 8 percent. Compared to 1950, people today are living much longer, are retiring at an earlier age, and are having fewer children. The results of these trends are presented in Table 1. As shown in the table, there are now more than 21 million people age 65 or over, comprising more than 10 percent of the U.S. population. Furthermore, the senior citizen market is growing over twice as fast as the rest of the population.

Even though these trends have been going on for many years, there is surprisingly little known about the consumption behavior of this market. As one source explains it, "The American obsession with youth has caused most businessmen simply to ignore the elderly." (Forbes, 1969). Indeed, from 1950 through 1970 the youth market was growing much faster than the senior citizen market, and the average age of the population dropped by more than two years. This trend has now been reversed, and the average age of the population will rise in the next few years, paralleling the increasing importance of the senior citizen market.

At the 1970 White House Conference on Aging, it was reported that the aggregate income of persons age 65 and over was approximately \$65 billion in 1967, more than double the level of 1958 (Chen, 1971). The 1970 Census indicated this figure had grown to \$68 billion. Expenditures by senior citizens was estimated in 1971 at \$60 billion, "far larger than the vaunted and highly amorphous youth market -- variously estimated at anywhere from \$20 to \$45 billion" (Business Week, 1971). In addition to being economically important, the senior citizen market is politically important. Comprising 10 percent of the total population, persons over age 65 account for 17 percent of the population old enough to vote, and they vote in much higher proportions, 70 percent, than younger voters (Business Week, 1971).

The average man age 65 today can look forward to thirteen more years of life, and the average woman to sixteen more years. By the end of the century individuals will be spending one quarter of their lives in retirement (Forbes, 1969). Because of these trends, it is important for consumer behavior researchers to begin to study the attitudes and behavior of this increasingly important market segment.

Most of the research that has been done investigating the senior citizen market has focused on the implications of the increasing numbers of individuals age 65 and over, their share of aggregate expenditures, and the type and sizes of products purchased by this group (Reinecke, 1964; Goldstein, 1968; Media Decisions, 1973; Forbes, 1969; Business Week, 1971). For example, Reinecke reported that the differences in general patterns of expenditure of older consumers were due almost entirely to a lower income and smaller number of individuals per household (Reinecke, 1971). Age itself had little influence, with the exception that medical and utilities expenditures rose with age and clothing expenditures tended to decline with age.

Although Zelan (Zelan, 1969) reports that interviewing the aged is not a problem, very little research has been conducted on consumer behavior patterns of the senior citizens. Schiffman has studied perceived risk (Schiffman, 1972), the importance of informal sources of information (Schiffman, 1971), and social interaction patterns (Schiffman, 1972) involved in the purchase decision for a salt substitute product. Mason and Smith conducted an exploratory study on the shopping behavior, distance traveled, and sources of information used by elderly residents of a public housing project (Mason, et.al., 1973). Salmi and Palubinskas studied the expenditures, leisure activities, and media habits of middle and upper class attendants of a senior citizens' center.

The major limitation of each of these studies is the population used for the investigation. Each used a captive audience, residents of a single building (Mason et.al., 1973; Schiffman, 1971 and 1972) or attendants of a single senior citizen facility (Salmi, et.al., 1972). In addition, the sample sizes used in these studies are quite small, over 100 only in the Salmi and Palubinskas study in which 211 (55 percent) of the attendants to the center completed the self administered questionnaire. There is thus a need for a study of the consumer behavior of the elderly using a population and sampling scheme which allows the examination of a more representative group of senior citizens.

The purposes of this study is to profile the senior citizen market, using a large probability sample of the population of a major metropolitan area. Specifically, this study investigates the (1) personal characteristics, (2) shopping behavior, (3) credit usage, (4) media habits, and (5) leisure time activities of the senior citizen market.

<sup>1</sup>The authors wish to thank Ferguson Rood and Edward Mumford of the Atlanta Journal and Constitution for permission to use the data in this paper and for their valuable assistance on this project.

## The Study

### Methodology

Personal in-home interviews were conducted with a proba-



TABLE 1  
Trend in Population  
of U.S. Senior Citizens

	1950	1960	1970	1973
Number of people age 65 or older (thousands)	12,300	16,560	20,065	21,300
Percent of total population	8.1%	9.2%	9.9%	10.2%
Average annual percent change over previous period, total population	---	1.7%	1.3%	.9%
Average annual percent change over previous period, those age 65 or older	---	3.0%	1.9%	2.0%
Median age of population (years)	30.2	29.5	28.0	28.4

Source: U.S. Bureau of Census, U.S. Census of Population: 1950, Vol. I, 1960, Vol. I; 1970, Vol. I, Part B; and Current Population Reports, Series P-25, No. 519.

TABLE 2  
Personal Characteristics of Senior  
Citizens vs. Other Age Groups

	18-34	AGE 35-49	50-64	65 or Older
<u>Sex</u>				
Male	48%	49%	47%	38%
Female	52	51	53	62
<u>Income</u>				
Under \$5,000	14	13	25	61
\$5,000-9,999	33	27	33	22
\$10,000-14,999	26	24	16	8
\$15,000 or over	27	36	26	9
<u>Race</u>				
White	80	79	78	78
Nonwhite	20	21	22	22
<u>Education</u>				
College Graduate	16	19	15	9
Part College	24	15	13	9
High School Graduate	35	37	30	17
Part High School or Less	24	29	42	65
n =	(1357)	(955)	(735)	(380)
Weighted n =	(2995)	(1928)	(1329)	(739)

TABLE 3  
Department Store Shopping Behavior

Department Stores	AGE				Total
	34 or younger	35-49	50-64	65 or Older	
Arlan's					
Heavy	5.0%	3.8%	1.6%	0.0%	3.5%
Light	6.0	2.0	4.6	3.8	4.4
Davison's					
Heavy	8.1	10.9	10.2	10.0	9.4
Light	19.6	19.0	20.6	12.3	18.8
G.E.X.					
Heavy	3.8	2.2	1.7	1.4	2.7
Light	9.0	7.5	5.3	2.7	7.2
J.C. Penney					
Heavy	13.5	15.0	8.8	5.1	12.1
Light	26.8	26.4	15.5	10.2	22.8
K-Mart					
Heavy	16.0	19.0	19.2	5.3	15.3
Light	26.3	24.6	19.4	18.6	23.7
Rich's					
Heavy	20.0	20.5	20.1	16.0	23.0
Light	28.5	29.7	26.7	23.8	27.9
Richway					
Heavy	12.0	8.8	4.9	2.9	8.8
Light	22.8	20.9	15.6	6.6	19.2
Weighted n	(3000)	(1942)	(1352)	(735)	(7035)

TABLE 4  
Credit Card Ownership

Card Ownership	AGE				Total
	34 or younger	35-49	50-64	65 or Older	
Possesses store credit card	32.6%	41.4%	38.4%	16.6%	34.4%
Possesses gaso-line credit card	37.2%	47.5%	45.2%	23.0%	40.0%
Weighted n	(2955)	(1928)	(1329)	(739)	(6959)

bility sample of 3,435 male and female adults age 18 and older in the 15 county Atlanta, Georgia SMSA. Eleven percent of these interviewed, 380 people, were age 65 or greater. The results were weighted to balance properly for individuals difficult to find at home, number of adults at home at the time of interview, day of the week the interview was conducted, and several population characteristics. The interviews were conducted on evenings and weekends between November 1972 and July 1973 by Belden Associates of Dallas, Texas as part of a major study for the Atlanta Journal and The Atlanta Constitution, Atlanta's major newspapers.

#### Results: Personal Characteristics

The personal characteristics of senior citizens compared to other age groups was investigated first, and the results are presented in Table 2. Because the life expectancy of females is greater than that of males, females comprise 62 percent of the over 65 age group compared to slightly over half for the other age groups. Most of the senior citizens were retired, 64 percent, resulting in a significantly lower income distribution than the other age groups. Well over half of the elderly earned less than \$5,000 per year, versus only one quarter or less of the younger age groups. In addition, the proportion of younger households earning \$15,000 or more was three to four times as great as the household headed by senior citizens.

The proportion of white, blacks, and other racial groups among the senior citizens was approximately the same as that for other age groups. As might be expected, there is a large difference in the education level of senior citizens compared to other age groups. Almost two thirds of those age 65 and older have not completed high school, and only 9 percent are college graduates. The extent to which these differences in personal characteristics, higher percentage of females and lower income and education, result in differences in consumer behavior will now be examined.

#### Shopping Behavior

Table III presents department store shopping behavior by age category for a number of different stores. Separate data are presented for heavy versus light usage, with three or more trips in the previous month defined as heavy usage. Senior citizens are significantly less likely (at the .01 level) to use discount stores like Arlans, GEX, K-Mart, or Richway, than all other age groups. For example, only 5 percent of the elderly made three or more trips to K-Mart in the previous month, compared to between 14 and 19 percent of the other age groups.

While they are much less likely to shop at discount stores, it appears that senior citizens are a significant supporter of the traditional department stores in the city. The elderly appear willing to travel to the downtown shopping district and to pay higher prices for department store merchandise. Further research should be conducted to determine the cause of the senior citizens' preference for traditional department stores over the discount stores. The research should seek to determine if the differences described here are attributable to merchandise offerings, transportation problems, or image and attitude differences.

#### Credit Card Possession and Usage

Data on store and gasoline credit card ownership by age category is presented in Table IV. Only one senior citizen in six possesses a store credit card. The comparable figure for the other age categories is twice this

level or more. The pattern of ownership of gasoline credit cards is similar, with less than one quarter of the elderly owning a gasoline credit card, compared to between 37 and 48 percent of the younger citizens. Unfortunately, this research did not seek to determine the cause of these differences (which are significant at the .01 level). It would be interesting to find out why so few senior citizens have a store credit card, even though they are as likely to shop the traditional department stores as the other age groups.

#### Media Habits

In the study, detailed media habit information was collected. Space constraints here require that only a verbal summary description be presented. The findings are:

- (1) Over half of the elderly read a daily newspaper, and almost 70 percent read the Sunday newspaper. This readership level is higher than that of the under-35 age group, and is somewhat less than the level of the 35-64 age group.
- (2) The senior citizens are slightly less likely than other groups to be readers of home-oriented magazines, such as Good Housekeeping and Ladies Home Journal, and of general readership magazines such as Reader's Digest. They are very much less likely than the other age groups to read the news weeklies (Time, Newsweek, U.S. News and World Report), and also magazines such as Playboy. It appears that the elderly read selectively and prefer to get their news from newspapers and broadcast media.
- (3) They listen to A.M. radio as much as others, but are much less likely to listen to F.M. Their listening is primarily at home, with only 19 percent listening to the radio in the car, compared to 48 to 64 percent for the other age groups. The heaviest listening period for the elderly is between 6:00 a.m. and 9:00 a.m., when about 25 percent of them listen to radio, about the same level as the other age categories. During the afternoon and evening hours, the listening rates fall dramatically. For example, only about 6 percent are listening in the 6:00 p.m. to 7:00 p.m. time slot, versus double that level for the other age categories.
- (4) The elderly are heavy watchers of early morning television, with over one quarter of them watching television before 9:00 a.m., twice as many as the other age groups. From 9:00 a.m. until 6:00 p.m. the senior citizens are also heavy TV watchers, with a higher percentage watching (approximately 33 percent) than the other age categories. The elderly appear to be television news oriented, as indicated by the fact that the period between 6:00 p.m. and 7:00 p.m. draws about 66 percent of all the elderly, by far the highest percentage of an any age group. After 9:00 p.m. their rate of watching begins to fall, and after 10:00 p.m. they are much less likely to watch than the other age groups.

#### Leisure Time Activities

The study measured respondent involvement in over seventy leisure time activities. Only the most interesting findings in relation to the elderly will be presented here. As a group the elderly spend their time reading books (43%), sewing (27%), attending baseball games (15%), fishing (15%), going to movies (11%), doing crossword puzzles (10%), gardening (31%), attending church activities (58%), going out to eat (51%). They are also heavy users of long distance telephone (65%), taxi cabs (32%), mass transit (43%), and laundromats (23%). These are the activities where they spend their time. However, in almost all instances even for these activities, they do them less than the other groups.

Their pattern of activity is similar to other groups, but their activity level for each is lower.

The elderly do not take part in physically oriented activities very much, as we might expect. However, there are some activities that are open to them in which they have very low involvement. These include, nonvoting political activity (5%), attending fashion shows (7%) and attending sporting events other than baseball (average about 3%).

#### Conclusion

This article has presented a profile of the senior citizen market on a number of dimensions. The findings indicate that the market is large, economically important, has specific shopping, credit and media habits, and takes part in particular types of activities. It is the authors hope that this paper will stimulate others to examine this much neglected and under researched segment of our population.

#### References

- Yung-Ping Chen, "Income Background," Report of the White House Conference on Aging, Washington, D.C., 1971.
- "Don't Write off the Senior Citizen Markets," Media Decisions, (July 1973), 64-67; 112-116.
- "The Forgotten Generation," Forbes, (January 15, 1969), 22-29.
- Sidney Goldstein, "The Aged Segment of the Market, 1950 & 1960," Journal of Marketing, 32 (April 1968), 62-68.
- Joseph Barry Mason and Brooks E. Smith, "An Exploratory Note on the Shopping Behavior of the Low Income Senior Citizen," The Journal of Consumer Affairs, (Summer 1973), 204-210.
- "The Power of the Aging in the Marketplace," Business Week, (November 20, 1971), 52-58.
- John Reinecke, "The 'Older' Market -- Fact or Fiction?" Journal of Marketing, (January 1964), 60-64.
- A. Coskun Samli and Feliksas Palubinskas, "Some Lesser Known Aspects of the Senior Citizen Market -- A California Study," Akron Business and Economic Review, (Winter 1972), 47-55.
- Leon Schiffman, "Perceived Risk in New Product Trial by Elderly Consumers," Journal of Marketing Research, 9 (February 1972), 106-108.
- Leon Schiffman, "Social Interaction Patterns of the Elderly Consumer," AMA Proceedings, (1972), 445-451.
- Leon Schiffman, "Sources of Information for the Elderly," Journal of Advertising Research, 11 (October 1971), 33-37.
- Joseph Zelan, "Interviewing the Aged," Public Opinion Quarterly, (Fall 1969), 420-424.

A TRANSGENERATIONAL COMPARISON ---  
THE ELDERLY FASHION CONSUMER

Claude R. Martin, Jr.  
The University of Michigan

Abstract

This paper reports the characteristics of elderly women fashion consumers and then compares their shopping and buying behavior with other generations of women consumers. The conclusion is that major differences exist among generations with particular emphasis on alternatives considered, predispositions formed prior to shopping embarkation, and reliance on sales persons and newspapers.

Introduction

"They feel slighted and ignored as consumers" (Business Week, 1971).

That assessment of the elderly consumer is particularly true in the case of women's fashion apparel based on three observations:

1. The consumer behavior literature does not contain reports on research into the buying needs or behavior of this market.
2. In-depth interviews with 75 women consumers, including 18 past the age of 60, confirmed John Howard's observance that "you have a psychological thing here, many older people don't want to be reminded that they are old, and they often tend to react against advertising and marketing programs that separate them from the masses" (Business Week, 1971).
3. A perusal of fashion advertising shows an emphasis toward the "Pepsi generation", or at best, the mothers of that generation.

This emphasis toward the youth and mid-age market continues despite an over-65 market of 20-million persons with more than \$60 billion spending power. The commonly cited statistic that women out-live the male of the species lends credence to the notion that the elderly woman is a viable market segment for the fashion apparel industry.

This article presents the results of a study designed to examine transgenerational differences and similarities in buying behavior, with emphasis on the geriatric fashion buyer.

The Study

The ultimate research instrument was a mail questionnaire, preceded by the previously cited series of in-depth interviews. These interviews with women consumers were conducted in southwestern Missouri. Each of the 75 women were approached just after completing an apparel purchase. They were asked to retrace the steps they had taken in consummating that purchase and the results of these in-depth interviews were used to structure the mail questionnaire.

Two Missouri retail trade areas, Joplin and Springfield were used for the mail survey. Among the rationale for that selection were that (1) the retailers agreed to cooperate, and (2) the two areas had substantial differences in socioeconomic characteristics (U.S. Census, 1970). Among these major differences were median and mean income levels, median housing values, population growth, and education level attainment. As a simple characterization, the Springfield market is larger, faster growing, better educated and more affluent than Joplin. Also important to the selection was the geographic closeness of the two cities; allowing for control of major regional variance.

The cross-index directories of the R.F. Polk Company were used to randomly choose 850 households, with the population proportion between the two markets as a factor in the sample selection. A telephone inquiry was conducted to obtain the name of women residents. These names were then numbered and one name in each household was randomly selected for receipt of a mail questionnaire. There were 356 completed questionnaires returned, with a reasonable proportionality between the two markets with regard to the age, population, marital status and employment distributions of the 1970 census.

The original in-depth interviews resulted in a hypothesis that there are transgenerational differences in buying behavior. Those interviews produced a series of specific variables for further exploration. In the subsequent mail survey each woman was questioned concerning her most recent personal clothing purchase. The 35 specific variables tested included those categorized as behavior, predispositions, product and information uses, buyer goals, and demographics (Figure 1). On this last dimension -- demographics -- the respondents were grouped into simple life-cycle classifications utilizing age and marital status. Of the 356 respondents, the following was the distribution:

	<u>Per Cent</u>
Single, 17 years or less	5.5
Single, 18 to 30 years	9.9
Married, 18 to 30 years	12.3
All respondents, 31-39 years	11.1
All respondents, 40-59 years	36.4
All respondents, 60 years and over	24.7

The ultimate objective of the study was to examine the multidimensional differences among various generations in these groups.

Typologizing the Elderly Consumer

However, before that trans-generation comparison an analysis of the simple cross-tabulation data on the

elderly woman was undertaken. The objective of this initial phase was to construct a typology by emphasizing those variables which typified the majority of the older consumers. The data reveal 25 dimensions which can be so used and to make the typology more meaningful these were rank-ordered in terms of percentage of elderly fitting the category. As seen in the results (Figure 2), the major factors that describe the elderly fashion consumer are that she buys a garment to match accessories already in her wardrobe, in a store where she has previously purchased; with no lower price limit coming to the market, but with a negative color predisposition -- there is at least one color she would not buy; and she utilizes newspaper advertising as an input to her decision process, but rarely a "shopping pal" -- she is a "loner". While the results do spell out other dimensions of this typology, perhaps some emphasis can be given those that coincide with very controllable strategies a marketer could evoke to sell this elderly consumer. Apparently she relies upon the salesclerk for point-of-purchase advice on both style and fit and upon newspaper advertising otherwise; she has definite fabric and garment care predispositions; and she also comes to the marketplace with a firmly fixed upper price limit. Perhaps most important overall is her self-perception that she keeps up to date on fashion trends. This latter finding that two-thirds of the elderly women exhibit a positive degree of fashion consciousness, coupled with an almost like degree of specific pre-purchase planning seems to explode the myth that the older woman is an exploitable target for left-over styles or inventory.

#### Transgenerational Comparison

##### MCA Analysis

The multiple classification analysis program (Andrews, et al, 1967) was used to detect those variables from among all tested, that had the highest relationship or best discriminated the life cycle variable. In its elementary form the MCA program produces measures of simple associations -- pairwise correlations -- between the dependent variable (in our case, simple life-cycle configurations) and the independent variable. This is reflected in the  $\text{Eta}^2$  score output of the program, which is the fraction of the variance in the dependent variable explained by the single variable alone. However, marketing does not function in a univariate environment and more meaningful for guidance was the  $\text{Beta}^2$  output of the MCA program.  $\text{Beta}^2$  is somewhat analogous to  $\text{Eta}^2$  but it is the measure of importance of the variable in a multiple regression of all variables to the dependent variable. Thirty four variables were used as independent variables with life cycle the dependent variable in the MCA analysis.

Arbitrarily, those 11 variables with a  $\text{Beta}^2$  score  $> .02$  were selected for further study (Table 1).

##### Generational Dichotomies

The small percentage of respondents in some age and marital categories caused the elimination or combining into the following groups:

- Post WWII adults (18 - 30 years)
- Matrons (40 - 59 years)
- Elderly (over 60 years)

The rationale for this selection was the heavy concentration of promotion and product mix toward the 18-30 year old woman. This is the generation being depicted by the psychographers as being very different from the previous passers in the life stream. Certainly a com-

FIGURE 1  
Buyer Construct

##### Demographics:

- Marital status
- Age
- Employment status of respondent
- Employment status of husband of respondent, if married
- Number and ages of children
- City of residence

##### Predispositions:

- Negative colors -- garment colors respondent would not buy
- Negative fabric characteristics -- fabrics respondent would not buy
- Garment care characteristics wanted
- Wardrobe accessory matching
- Upper and lower price limits to purchase
- Had charge account where shopping and buying reported
- Previously bought apparel in store of purchase
- Prepurchase planning:
  - General
  - Specific -- positive color wanted
  - positive fabric wanted

##### Product and Information Uses:

- Comparison shopping at alternate stores
- Utilization of price limitations
- Method of payment
- Sought out particular sales clerk
- Use of "shopping pals"
- Used sales clerk evaluations of style and fit of garment
- Evaluation of mass media helpfulness in purchase decision

##### Buyer's Goals:

- Self-evaluation of fashion awareness
- Factors used in developing level of fashion awareness
- Shopping enjoyment in buying clothes for self

##### Behavior:

- Coordinating items purchased
- Type of garment purchased
- Number of stores shopped
- Number of stores shopped on day of purchase
- Color of garment purchased
- Fabric of garment purchased
- Garment care requirement for item purchased

FIGURE 2  
Elderly Fashion Consumer Typology

Characteristic	Per cent pf Elderly Women
The garment purchased matched accessories now in wardrobe	91.5
Woman had purchased clothes for herself in store before	88.2
Did not come to the market with a lower price limit	83.3
Before shopping had formulated a negative color predisposition	80.5

Figure 2 -- continued

<u>Characteristic</u>	<u>Per cent of Elderly Women</u>
Used newspaper advertising at least fairly often for guidance in buying clothing	80.3
Seldom or never shop with other persons	80.3
Sales clerk's evaluation of garment fit was helpful in purchase decisions	80.0
Sales clerk's evaluation of garment style was helpful in purchase decisions	77.9
Newspaper advertising is helpful in buying own clothes	74.1
Preference for permanent press fabric treatment	73.4
Would not buy a garment that had to be dry-cleaned	73.1
Did not specifically seek out the sales clerk from whom purchase was made	68.8
Before shopping had formulated a specific fabric preference	67.1
Maintained a charge account at the store of purchase	67.1
Evaluate self as keeping up-to-date on fashion trends	66.7
Before shopping had formulated an upper price limit and stayed under that limit in actual purchase	63.4
Would have made personal clothing purchase from a male clerk	63.0
Before shopping had formulated specific plans concerning purchase	61.3
Shopped in a total of 2 or less stores before making purchase	59.0
Did not have a positive color preference before beginning to shop	58.2
Evaluate self as enjoying shopping for clothes very much	55.0
Shopped in only one store on day of purchase	54.1
Shopped in a high-fashion, women's specialty store	53.7
Bought garment by paying cash	52.5
Made purchase in high-fashion women's specialty store	50.0

parison of those in the next generation down from the elderly also offers the possibility for filling out any indicators of the pros and cons of transgenerational marketing strategies. In other words, these three age groupings provide a basis for exploring whether

marketers can bridge across the spectrum of women fashion consumers.

TABLE 1  
MCA Analysis - Variables with Beta<sup>2</sup> > .02

<u>Variable</u>	<u>Beta<sup>2</sup></u>
Frequency of use of newspaper advertising	.072
Sales clerk help on style and fit <sup>1/</sup>	.040
Purchase made in a specialty boutique store <sup>2/</sup>	.036
Item purchased	.035
Shopped in specialty boutique store before making purchase <sup>2/</sup>	.031
Shopping enjoyment	.030
Newspaper advertising helpful in purchase decision	.030
Number of stores shopped on day of purchase	.029
Higher price limit before shopping	.024
Self-evaluation of fashion awareness	.024
Total number of stores shopped	.020

<sup>1/</sup> Because of the similarity across all the data on the evaluation of sales clerk help on style and fit, these two variables were merged into one single variable.

<sup>2/</sup> A five-member retailer panel was used to categorize the 96 stores mentioned by respondents into general store classifications.

A comparison of the elderly woman to those in the other two generations was made along the eleven dimensions (Table 2). For the young woman the greatest diversity with the elderly was on the garment purchased and the smallest difference was in the shopping enjoyment experienced. Similarly the greatest difference between the elderly woman consumer and her matron counterpart is in the number of total stores shopped, while the least difference is in a predisposed upper price limit. By rank-ordering these differences and using the Spearman rank correlation analysis (Siegel, 1956), the  $r$  is +.145-leading to a conclusion that major trans-generational differences do exist.

Examination of the results for each group (Table 3) gives a definition of the differences. For example, when we contrast the elderly woman with the younger one we see the greater proportion using and finding helpful both newspaper advertising and sales clerk evaluations in the purchase decision

We also can see the greater proportion of younger women enjoying shopping and entering the market place with an upper price limit. However, examination of the data on the latter dimension shows the following proportion of

TABLE 2  
TRANSGENERATIONAL DIFFERENCES  
(Percentage of Respondents)

TABLE 3  
TRANSGENERATIONAL CHARACTERISTICS  
(Percentage of Respondents)

Variable	Percentage Differences	
	Elderly vs Young	Elderly vs Matron
Used newspaper advertising often for buying guidance	19.5	10.8
Sales clerk helpful on style and fit	40.4	8.9
Bought garment in specialty boutique	17.7	3.1
Bought a dress	45.8	14.7
Shopped for garment in a specialty boutique	20.4	4.9
Enjoy shopping for own clothes	12.3	9.7
Shopped in a total of two or fewer stores	36.6	22.2
Mass media provided helpful purchase information	24.0	9.6
Shopped one store on day of purchase	17.6	13.2
Upper price limit	24.8	1.9
Keep up to date on fashion	25.2	8.8

Variable	Young Woman	Matron	Elderly
	Used newspaper advertising often for buying guidance	60.8	69.5
Sales clerk helpful on style and fit	38.5	70.0	78.9
Bought garment in specialty boutique	18.9	4.3	1.2
Bought a dress	17.6	48.7	63.4
Shopped for garment in a specialty boutique	21.6	6.1	1.2
Enjoy shopping for own clothes	98.6	76.6	86.3
Shopped in a total of two or fewer stores	32.4	46.8	69.0
Mass media provided helpful purchase information	50.0	64.4	74.0
Shopped only one store on day of purchase	33.8	38.2	51.4
Upper price limit	91.9	69.0	67.1
Keep up to date on fashion	91.9	75.5	66.7

each age group who adhered to an upper price limit:

Young	80.8%
Matron	92.6%
Elderly	94.5%

Classification Analysis (Ann Arbor: Institute for Social Research, University of Michigan, 1967).

Sidney Siegel, Nonparametric Statistics for the Behavioral Science (New York: McGraw-Hill, 1956).

The elderly woman shops in fewer stores totally and on the day of purchase than does either of her younger counterparts. The elderly also self-evaluate themselves more conservatively as to fashion awareness than the 18-30 year old, but similar to the matron.

#### Conclusion

The two-thirds of elderly women who perceive themselves as being fashion conscious and the high proportion who enjoy shopping seems to indicate a potential for fashion merchandising. Certainly they are different from the young woman who is the target of so many fashion retailers. These differences and the dichotomies over all three generations certainly indicate the difficulties of mounting a transgenerational strategic effort.

#### References

Business Week, November 20, 1971, pp. 52-59.

United States, Bureau of the Census. Census of the Population, 1970, PC(1)-A27, PC(1)-B27, PC(1)-C27. Washington, D.C.: Government Printing Office, 1971.

Frank Andrews, James Morgan, and John Sonquist, Multiple

A DESCRIPTIVE MODEL OF CONSUMER CHOICE PROCESSES  
AMONG NURSING HOME PATIENTS

Steven A. Baumgarten, Purdue University  
Tanniru R. Rao, University of Wisconsin-Milwaukee  
L. Winston Ring, University of Wisconsin-Milwaukee

Abstract

This article presents a model describing the processes by which consumers first select and then adjust to nursing homes. The model is based upon data collected through survey research, and it bears implications for both improvement of the consumer decision process, and reorientation of nursing home administration.

Introduction

In the period 1963-1970, the number of nursing homes in the United States increased from approximately 13,000 to over 20,000. During that time period, the number of patients in nursing homes almost doubled. Today, there are over one million patients in U.S. nursing homes. This represents approximately 4% of all persons over the age of 65.

Despite the obvious size and growth of Nursing Homes as an important societal institution, the process by which consumers select and adjust to nursing homes has not been the subject of systematic empirical investigation.

Viewing the relationship between nursing homes and their patients as a marketing system provides a useful conceptual framework for understanding the pre- and post-adoption decision processes of the elderly person faced with the decision of whether to enter a nursing home.

This article presents a descriptive model of consumer choice processes among nursing home patients which is based upon an empirical study sponsored by the U.S. Department of Health, Education and Welfare.

The Data

The data upon which this descriptive model is based were obtained from a wide variety of sources. Among these were:

1. Discussion with twenty nursing home administrators, social workers and therapists;
2. Individual depth and structured interviews with eleven elderly persons living in private residences, twelve elderly persons in hospitals, two relatives of nursing home patients, three doctors of nursing home patients, and one lawyer of a nursing home patient;
3. Group depth interviews with a total of eighty patients in three separate nursing homes;
4. Structured field interviews with 121 patients at sixteen area nursing homes, and
5. Structured interviews with floor nurses at each of the sixteen nursing homes.

The data were collected in late 1971 in the Milwaukee metropolitan area.

The Model

The model was generated strictly from the data and in-

formation collected in the research project, and is intended to be an accurate description of the process by which consumers actually do select and adjust to nursing homes.

The model in Figure 1 presents the process by which consumers select a nursing home. Figure 2 is, in effect, a continuation of Figure 1, and presents the process by which consumers adjust to the nursing home environment.

In both figures, the center column presents the basic model and shows the selection and adjustment processes sequentially, primarily as seen from the (potential) patient's point of view. The left- and right-hand columns contain empirically determined breakdowns of the components of the processes. (To avoid unnecessary verbiage, the elderly person involved in both the selection and adjustment processes will hereafter be referred to as "the patient," with the understanding that he or she is, in fact, only a potential patient until admission to the nursing home.)

Stimulus and Problem Identification

The model begins with a stimulus leading to identification of the problem, "Should a nursing home be entered?" In almost all cases, the stimulus seems internal--self-recognition of the problems of old age, illness, inability to care for oneself, etc. Only 4% of the nursing home patients surveyed indicated that a doctor's recommendation was a major stimulus, and only 6% expressed the stimulus altruistically as a desire not to burden others. The insistence of relatives was a stimulus for only 4% of those surveyed, but as is seen later, relatives' input and/or participation in the decision-making process is a very important element in the model.

Value Conflict

Confrontation of the problem, "should a nursing home be entered?" almost invariably leads to conflict within both the patient and the family, and between the patient and the family.

The family frequently has guilt feelings over "abandoning" a loved one to the care of an impersonal institution. This guilt is generated both internally, by the felt need to be personally responsible for the care of a loved one, and externally by residual social mores which dictate that people should be personally responsible for the care of loved ones.

The patient desires family togetherness, a sense of belonging and frequent love and attention, as well as peer relationships and (often) more extensive care than can ordinarily be offered by family or friends. At the same time, the physical act (or even the very thought) of entering a nursing home can serve to emphasize the relative imminence of death.

Furthermore, both the patient and the family are very likely to have a negative image of nursing homes generated by unfavorable publicity or interpersonally-received "horror stories."



## Alternative Search

Family and patient conflict usually leads to a prolonged and frequently exhaustive search for non-nursing home alternatives. Living with or near relatives or friends, other types of care facilities ("foster homes," retirement homes, etc.), periodic extended "visits" with a number of family members, and paid companions or nurses are alternatives which are usually considered at this stage.

Sometimes an alternative solution is found and is, in fact, desirable. In many instances, alternative solutions are tried even when both the patient and family recognize that they are "logically" inferior to the nursing home solution. In both of these instances, the problem of entering a nursing home is likely to arise again in the future.

## Conflict Resolution

In most cases, conflict is not actually resolved, or even appreciably abated. Tension is reduced somewhat due to the relative comfort of finally being able to make a decision, but guilt feelings remain and a very high level of dissonance is generated among family members. The patient often feels abandoned and remains apprehensive about entering a nursing home.

Frequently, too, doctors, social workers and visits to nursing homes and other care facilities may help to reduce tension and, on a post-decision basis, dissonance.

## Decision

When the decision is one of entering a nursing home, it is very frequently made by persons other than the ultimate recipient of the services--51% of the nursing home patients indicated that the decision that they would enter a nursing home was made strictly by others. At the same time, only 24% claimed to have made the decision solely by themselves.

## Development of Choice Criteria

Once the decision to enter a nursing home is made, selection criteria are usually necessary in order to priorly reduce the number of alternatives to a manageable size. With 121 nursing homes in the Milwaukee area, it would seem necessary to pre-screen on some basis. Despite the seemingly large number of nursing homes available, however, the nature of their offerings is so segmented that for most families, very few homes are suitable. Many homes are affiliated with specific religions and/or church-supported. There are small, medium and large homes; old and new homes; Medicare approved homes, Medical Assistance approved homes and homes with no assistance approval; Skilled care homes, limited care homes and personal care homes, etc.

## Generation of Alternative Set

The segmentation of these offerings, taken together with patient and family needs concerning location, facilities, level of care, religious affiliation, reputation, size, etc., typically results in the generation of a highly limited alternative set. In fully 78% of the cases, only one home was considered.

## Evaluation of Selected Alternatives

The bases upon which a nursing home is finally selected are also extremely limited. The most frequently mentioned criterion (21%) was "favorable impression of facilities." 13% relied upon the advice of others, 10% selected on the basis of religious affiliation and 7%

on the basis of location. Only 3% were concerned with medical care and only 2% with cost. In 36% of the cases, the patient indicated that he or she didn't choose the nursing home or didn't know what the criteria were. Thus the non-participative nature of the decision process continues through to the ultimate selection of the home.

## Selection to Admission

Now the nursing home has been selected. The patient may undergo a waiting period--typically less than one month, but in the case of some few of the most highly desired homes the waiting period may exceed one year--and is finally admitted to the nursing home.

As is the case with other consumer products and services, however, the consumer decision process does not end here. Post-purchase reactions, use of the product, etc., are important considerations for the marketer concerned with providing consumer satisfaction. In the case of nursing homes, the extent to which the patient successfully adjusts to the home is a primary determinant of the extent to which the "marketer" is successful.

Figure 2 presents a model of the patient adjustment process which is discussed in the following sections.

## Admission--Initial Shock

The initial reaction of a new patient to a nursing home can be, and usually is, traumatic. Suddenly there are rules (many of which, fostered by legislation, seem to have no logical basis), specified eating times, an almost total lack of privacy (very few patients have private rooms), an unfamiliar environment and the faces of only strangers.

## Institutional Adjustment Process

The process of adjusting to this new environment is only marginally aided by most institutions. Even a brief orientation program is more the exception than the rule. The patient, then, must rely upon other patients or friendly staff members to aid him or her in adjusting. In the absence of formal aids to adjustment, the new patient can also take solace in the improved medical security, improved care and increased number of planned activities now available to him or her.

## Internalization of a New Value System

In order to successfully adjust to the nursing home environment, it is often necessary for patients to revise their attitudes, interests and opinions. Thus, the patient who views nursing home life as a form of independence from being a burden to others can generally cope better than the patient who feels abandoned to an institution. Similarly, patients are found to be better adjusted when they develop the spirit of having a home within the nursing home; when, within their physical limitations, they can do useful things (read to fellow patients, write letters, tidy up, etc.); when they generate a social service orientation; and when they become involved, in almost any way, with events in the home and in the community.

The patient who successfully internalizes these values generally becomes more active, demands more from the home and the community, becomes politically involved, becomes a spokesman for the home, and spends little time worrying about death.

The typical profile of the patient who does not successfully internalize these values is that of passivity, a sense of resigned living, substantial bitterness, loneliness, feelings of rejection and high fear of death.

The successfully adjusted patient represents a societal gain. He or she is relatively satisfied and a contributing member of society. The unadjusted patient is unsatisfied, unhappy and a loss to society.

#### Implications

If the model presented above is viewed as one involving the relationship of a marketer and his customers, then two categories of implications may be meaningful: 1) Implications for improved consumer decision-making, and 2) implications for improved marketing and management of nursing homes.

##### Improved Consumer Decision-Making

One of the more disturbing features of the model is the general lack of patient participation in the decision-making process. It is interesting that our society has a great appreciation for the value of participative decision-making in family relationships, personnel management and school situations, to name just a few. Yet, in the decision to radically alter the life style of an elderly person, that person is denied participation more than half the time! (It should be noted that all of the 121 patients surveyed were lucid and capable of relatively sophisticated interaction.) If participation in decision-making does, in fact, lead to greater commitment to the decision, then certainly families considering nursing home care should be encouraged to involve the potential patient in the decision.

Once the decision to join a nursing home has been made, the vast majority do not, as one might expect, shop extensively. On the contrary, fully 78% of those surveyed considered no home other than the one they ultimately selected. We would be remiss if we didn't point out that the average teenager at a candy counter generates a more extensive alternative set!

In addition to this highly limited generation of alternatives, the evaluative criteria ultimately selected leave a great deal to be desired. "Favorable impression of facilities" (21%) and "Advice of others" (13%) represent the primary bases upon which the nursing home is selected. Returning to the consumer goods arena for another damning analogy, it is obvious that most people generate more thorough evaluative criteria for the purchase of a washing machine!

If the consumer is to make a reasoned choice in this very important decision process, he or she will have to first recognize the importance of the decision and then be provided the information necessary to more thoroughly evaluate the alternatives. What is needed is massive consumer education. Nursing home trade associations as well as individual nursing homes are in an excellent position to provide useful information to potential consumers. Additionally, Social Security Administration Offices and any other agencies which have frequent contact with the elderly are prime outlets for information dissemination and counseling.

##### Improved Management of Nursing Homes

One critical stage in the selection process is that of "Conflict Resolution." As pointed out earlier, conflict is usually not truly resolved at this stage, and it is here that nursing home administrators can begin to undertake marketing effort. It should be noted that nursing homes are frequently enjoined from competitive advertising by either industry association or legislative regulations. Most nursing homes and associations can, however, engage in generic promotion of the nursing home concept. Any form of reassuring and/or educational promotion is likely to reduce tension at the

conflict resolution stage and may well result in: 1) a more comfortable decision process for the patient and family, 2) a greater likelihood of the nursing home alternative being selected, and 3) a generally improved image of nursing homes among the population. It is probable that strong industry association activity will be necessary to accomplish this task.

If such promotion is undertaken, it is also likely to improve the subsequent stages of the selection decision process, particularly if the promotion is educational in nature (i.e., messages which explain nursing home procedures and provide a reasoned basis for consumer selection). Thus, evaluative criteria might be improved, thereby leading to better selection.

As an additional benefit, if consumers are more comfortable in their decision, they will likely be more positive in their attitudes toward the nursing home and likely to adjust more easily. Currently, 56% of the patients surveyed stated that they would have preferred not to come to a nursing home at all. It is perhaps not surprising that many fail to adjust successfully, given such a high negative predisposition.

In terms of the adjustment process, the initial shock of institutional life probably represents the greatest deterrent to rapid adjustment. The lack of privacy, lack of freedom, routinization and disruption of previous social activities are aspects of nursing home life which may be easily changed. Certainly, nursing home administrators should explore the possibilities of such changes as revised (more informal) scheduling of events and activities, providing of more private rooms, relaxing some non-legislated rules, and continuing (in some fashion) some of the previous social activities of the patients.

It is almost too obvious to point out that a simple orientation program and/or small-scale social event should accompany the admission of a new patient. Yet many nursing homes do little more than show a new patient his or her room. Certainly a welcoming "party" and/or orientation program would make the patient feel more "at home."

Finally, planned events and activities which encouraged interaction outside of the home could help to alleviate patient's feelings of isolation and uselessness. Instead of (or, in addition to) a bus trip to the zoo, some patients might benefit more from an opportunity to help work on a political campaign or charitable drive, by doing envelope stuffing, telephoning, etc. Similarly, events which bring the public to the nursing home, such as art shows and open houses, can also help alleviate feelings of isolation.

#### Conclusion

A better understanding of the processes by which patients select and adjust to nursing homes can provide insights into developing procedures for improved consumer decision-making. At the same time, it highlights areas in which administrative reorientation can provide increased customer satisfaction.

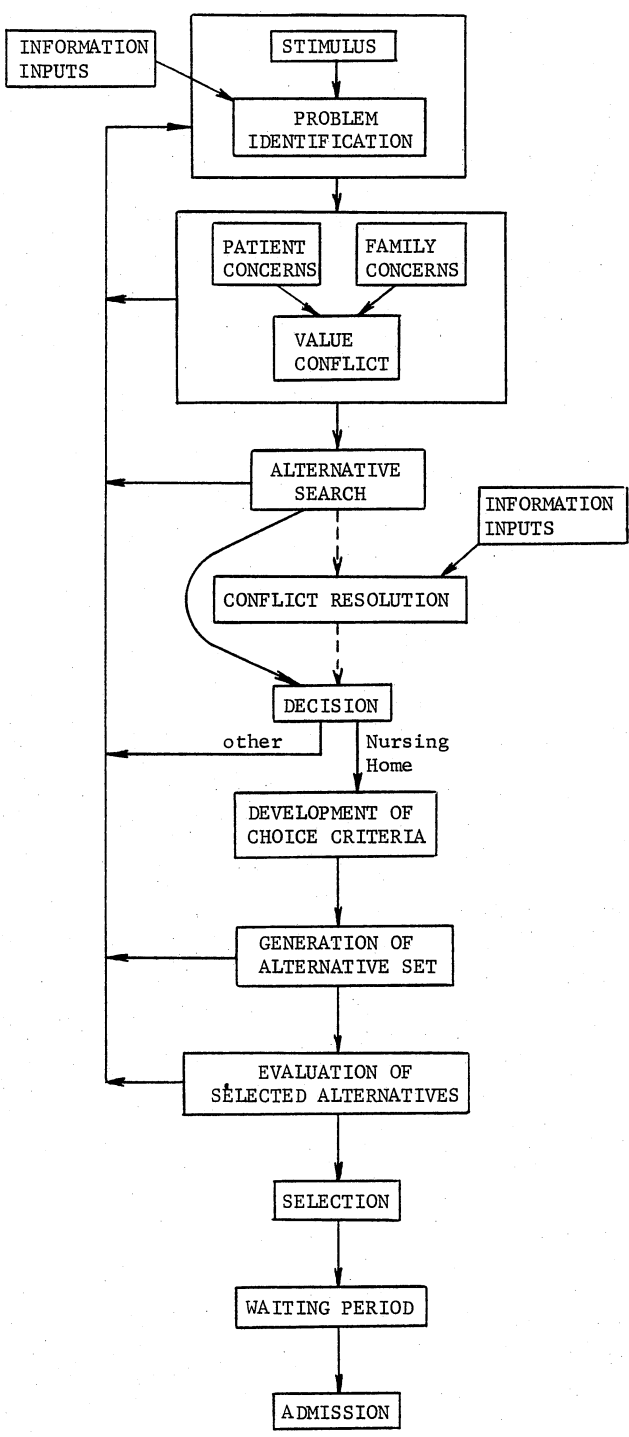
The presentation of survey research results in a descriptive model format is seen as desirable in that it: 1) Conceptualizes what might otherwise appear to be a loose collection of findings, and 2) helps to highlight empirically testable issues and thereby provides future researchers with a stronger conceptual framework for constructing data collection instruments and/or experimental procedures.

The model presented in this article focuses upon the

the decision-making process as seen from the patients point of view. This is a necessary first step toward understanding this complex process. It should be emphasize, however, that detailed investigation of the role of the patient's family in this decision process is a necessary next step for the development of a more comprehensive model.

FIGURE 1

doctor  
 Relative  
 Social Worker  
 Friends  
 -----  
 Should a nursing home  
 be considered?  
 -----  
 Stereotyped image of  
 Nursing Homes  
 Need for family  
 togetherness  
 Need for love and  
 attention  
 Fear of death  
 Need for peer  
 relationships  
 Need for medical care  
 -----  
 Join relative or friends  
 Live near family, friends  
 Other care facilities  
 Nursing Home  
 Status Quo  
 -----  
 Nice facilities  
 Location  
 Religion  
 -----  
 Favorable impression  
 of facilities . . . 21%  
 Advice of others . . . 13%  
 Religion . . . . . 10%  
 Location . . . . . 7%  
 Prior familiarity . . . 4%  
 Availability . . . . . 3%  
 Medical Care . . . . . 3%  
 Cost . . . . . 2%  
 Didn't choose . . . . 17%  
 Don't know . . . . . 19%



Illness/Old Age . . . 63%  
 Relatives' insistence. 4%  
 Doctor's recommendation 4%  
 Desire not to burden  
 others . . . . . 6%  
 Loneliness . . . . . 6%  
 Other . . . . . 17%

Stereotyped image of  
 Nursing Homes  
 Need for freedom  
 Need to care for loved ones  
 Social pressures  
 Guilt feelings  
 Patient benefits

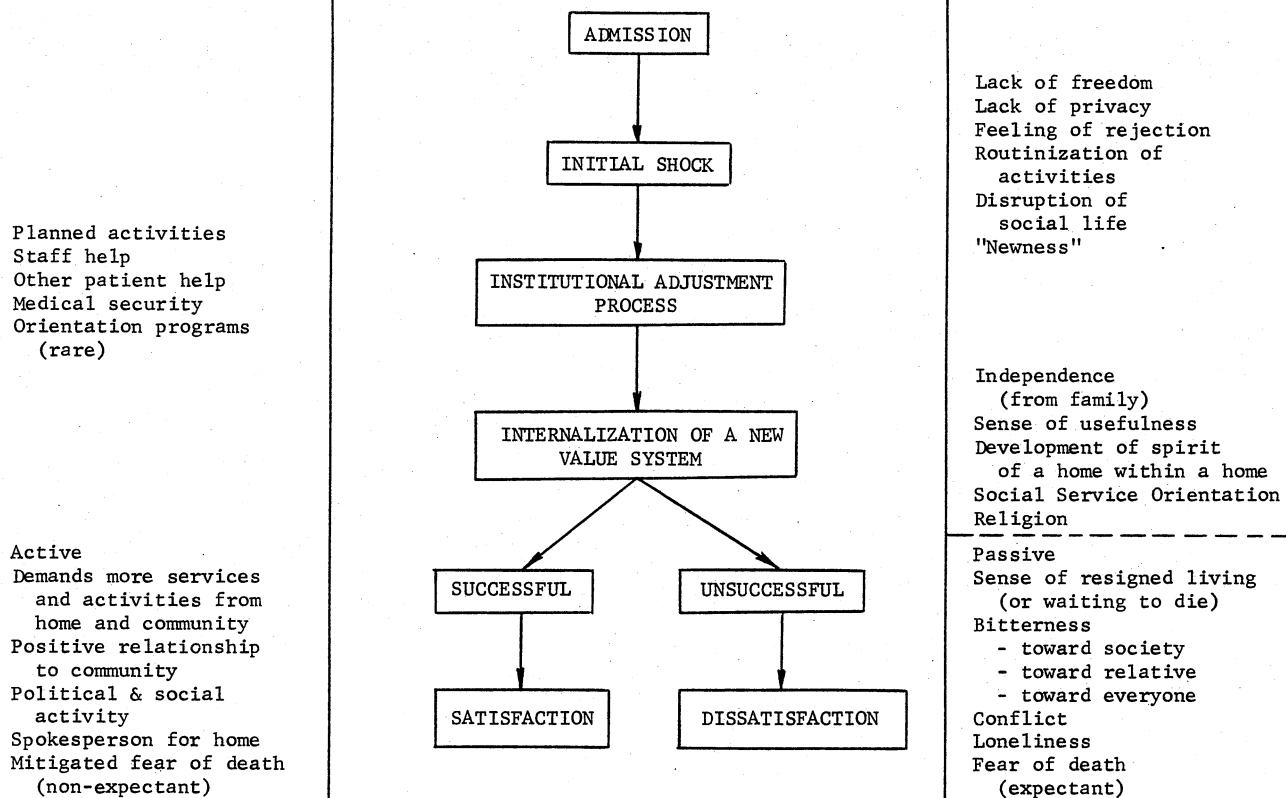
Doctor  
 Social Worker  
 Visits to Care Facilities

made by:  
 Self . . . . . 24%  
 Self with others . . . 11%  
 Joint . . . . . 14%  
 Others . . . . . 51%

Number of Other  
 Alternatives Considered  
 0 . . . . . 78%  
 1 . . . . . 7%  
 2 . . . . . 4%  
 3 . . . . . 5%  
 4+ . . . . . 5%

No wait . . . . . 60%  
 One month . . . . . 15%  
 2 months - 1 yr. . . . 3%  
 More than one year . . 22%

FIGURE 2



## THE ELDERLY CONSUMER: ONE SEGMENT OR MANY?

Jeffrey G. Towle  
and  
Claude R. Martin, Jr.  
The University of Michigan

### Abstract

This paper reports on the segmentation of a sample of elderly consumers selected from the 1973 National Target Group Index Study. The six segments obtained were defined by self concepts of buying style using cluster analysis and were described in terms of their psychographic characteristics.

### Introduction

The overall population growth of the past twenty-five years, coupled with rising incomes and the proliferation of consumption options in the market place has helped make it advantageous for marketers to recognize they are not dealing with a homogeneous mass of prospects, but with a diverse mixture of sub markets. Each of these submarkets or segments can be defined by a commonality of demographic, psychographic and/or behavioral characteristics. A review of the recent literature shows a sustained interest in segmentation and sub-segmentation which is still "based upon developments on the demand side of the market and representing a rational and more precise adjustment of product and marketing effort to consumer or user requirements" (Smith, 1956).

Given the general acceptance of segmentation as a useful and beneficial strategy to both producers and consumers (Haley, 1969), little work has been done sub-segmenting the elderly consumer group. The tendency of marketers is either to treat the elderly, over 65 consumers as a more or less homogeneous group, or to pay virtually no attention to them at all. This means that the more than 20-million elderly consumers are generally regarded as a single market segment, distinguishable from other segments on the basis of age, but homogeneous when treated alone, as a single sub-population. The limited attempts at sub-segmentation have been principally by suppliers of goods and services designed for the elderly and have concentrated on a few socio-economic factors, demographic characteristics and geographic location. For other marketers, grouping the elderly consumer together simplifies their conception of that market and simplifies their decision-making process. However, in line with segmentation theory it may also diminish their profit optimization with this overall class of consumers, (Frank, et al. 1972; Haley, 1969).

This paper reports on an attempt to sub-segment the over-65 market on demographic, psychographic and behavioral dimensions. The objective was to explore whether viable sub segments exist and to describe that diversity among the elderly so that marketers might recognize it and adapt to meeting the needs of parts of that market.

### Data Base

The elderly consumers studied were drawn from the 20,137 consumer data base collected by the Axiom Market Research Bureau (AMRB) as part of the national 1973 Target Group Index (TGI). The TGI is a compre-

hensive media and product survey of slightly over 20,000 adult consumers that also collected demographic, psychographic and buying style data on each respondent. The overall study consists of a probability sample of adults aged 18 and over in the continental United States. From that base we systematically selected every 10th respondent aged 65 years and over, arriving at a 10% sample consisting of 209 elderly respondents.

The variables used in our analysis included 13 demographic; 20 psychographic; and 10 buyer style measures (Figure 1). The demographic variables are standard and include most of those identified as major demographic segmentation variables (Kotler, 1972). The psychographics are a somewhat novel, but logical, approach by TGI.

Psychographic variables are those which describe the personality traits and attitudes of persons (Wells, 1974). The variables in the Target Group Index were developed by a pre-test survey conducted nationally by AMRB (TGI Report, 1972).

TGI uses twenty self-concept measures (adjective groups) which were designed to assess how the respondent views himself or herself as a person (Figure 1). These twenty measures are a distillation of 304 adjectives, which were selected from among all adjectives in the dictionary which could be used to describe people. A factor analysis of a pilot study using the original 304 adjectives produced a reduced set of adjectives which were then incorporated into the TGI survey.

The twenty psychographic dimensions are, therefore, basic ways in which people describe themselves. The list was systematically arrived at, although it does not depend on any particular psychological theory of personality except to the extent that a measure of "real-self" or "ideal self" is hoped for. Landon found a relatively high correlation between real and ideal self concept and noted that "the debate over which self concept is more important may be inconsequential" (Landon, 1974).

The data base also measures "buying style," and this is also based on a factor analytic reduction of several originally proposed measures.

For purposes of this research, the data consist of straightforward demographics, self evaluation of some psychographic dimensions (what they like as people); and self-evaluation of "buying style" (what they like as consumers). The data base was a representative sample of 209 elderly (over 65 years of age) consumers.

### Analytical Framework

The use of buyer behavior variables for segmentation has been more successful than utilization of personality traits or other possible dimensions of a psycho-segmentation (Rewoldt, et al, 1973). The authors agree with this evaluation and suggest that the viability of buyer behavior as a segmentation base was predicated on its

FIGURE 1  
Analytical Variables

Demographic:

Census Region  
 SMSA-Location within (rural, urban, suburban)  
 Socio-economic Classification  
 Type of Neighborhood  
 Type of Dwelling  
 Home Owned/Rented  
 Years at Present Address  
 Sex of Respondent  
 Race of Respondent  
 Respondent's Education  
 Respondent Employment Status  
 If Unemployed, Retired or Temporarily Unemployed  
 Respondents Type of Business (now or previously)

Psychographic:

Affectionate (Passionate, loving, romantic)  
 Amicable (Amiable, affable, benevolent)  
 Awkward (Absentminded, forgetful, careless)  
 Brave (Courageous, daring, adventuresome)  
 Broadminded (Openminded, liberal, tolerant)  
 Creative (Inventive, imaginative, artistic)  
 Dominating (Authoritarian, demanding, aggressive)  
 Efficient (Organized, diligent, thorough)  
 Egocentric (Vain, self-centered, narcissistic)  
 Frank (Straightforward, outspoken, candid)  
 Funny (Humorous, amusing, witty)  
 Intelligent (Smart, bright, well-informed)  
 Kind (Goodhearted, warmhearted, sincere)  
 Refined (Gracious, sophisticated, dignified)  
 Reserved (Conservative, quiet, conventional)  
 Self-Assured (Confident, self-sufficient, secure)  
 Sociable (Friendly, cheerful, likeable)  
 Stubborn (Hardhearted, headstrong, obstinate)  
 Tense (Nervous, highstrung, excitable)  
 Trustworthy (Competent, reliable, responsible)

Buying Style:

Brand Loyal (Those who e.g. look for the name of the manufacturer on the package)  
 Cautious (Those who e.g. do not buy unknown brands merely to save money)  
 Conformists (Those who prefer to buy things that friends or neighbors would approve of)  
 Ecologists (Those who believe e.g. that all products that pollute the environment should be banned)  
 Economy-Minded (Those who shop around a lot to take advantage of specials or bargains)  
 Experimenters (Those who feel they like to change brands often for the sake of variety and novelty)  
 Impulsive (Those who e.g. often buy an item on the spur of the moment)  
 Persuasible (Those who e.g. believe that advertising presents a true picture of the products of well-known manufacturers)  
 Planners (Those who e.g. generally plan their purchases in advance)  
 Style-Conscious (Those who try to keep abreast of changes in styles and fashions)

starting point for defining sub-segments in the market, with the 33 demographic and psychographic dimensions as potential describers of those behavioral segments. The basis for our analysis was to form natural segments, based on self-concepts of buying style, that used individual respondents as the basic analytical unit in a cross-sectional study. The purpose was to define purchasing behavior groups or segments and describe those groups in terms of the psychographic and demographic variables.

Verification of Data

While a reasonably precise process was followed by AMRB in constructing the Target Group Index, including definition of the psychographic and behavioral variables, the decision was made to undertake another verification of these dimensions.

Canonical analysis was used to check the redundancy (Cooley and Lohnes, 1971, pp. 12 and 170-72) of the information in each set of measures using the psychographic variables as one variable set and the buyer behavior variables as the other. "The statistical decision rule is to determine a linear combination of variables in each set so that correlation between the sets is maximized" (Cox and Enis, 1972). Such an analysis of sets of personality and behavioral variables was reported by Kernan in 1968. Further support for using canonical analysis in this way is discussed by Alpert and Peterson (1972). They also observe that canonical analysis can be applied effectively in conjunction with other multivariate techniques. The suggestion has been, as is done by the TGI construction of the 30 psychographic and behavioral variables, that one might factor analyze variables to remove multicollinearity within variable sets prior to canonical analysis. Our analysis showed a fairly high canonical correlation between the first pair of factors (.5427), the second pair (.4991), and the third pair (.4800).

There has been some discussion (Ginter, 1974; Bass and Wilkie, 1973) of the problem of univariate responders when scaling techniques are used. To check for this phenomenon in the TGI data, a histogram of the variability of response across the 30 behavioral and psychographic variables was constructed. That histogram showed 12 per cent of the respondents had zero variation and another 2.3 per cent who had a very low degree of response variation (their standard deviation around the mean response was less than .600). It was decided that such responders, who we named "univariate responders" should not be included in subsequent analyses since they provided no information about themselves from the information theoretic standpoint (zero variance = zero information).<sup>1</sup>

The strictly univariate responders were extracted from our data base and a distribution of their responses showed that all had answered "not sure" (the mid-point on the five-point scale) for all of the psychographic and behavioral dimensions.

<sup>1</sup> The information content of a message symbol is the negative of the logarithm of the probability that this symbol will be emitted from the source, e.g.,  $p=1.0$  for univariate responders,  $\log$  of 1.0 = 0.0 (Engels, 1971).

operationalizability. Based on this we decided that the 10 buyer behavior variables might form a reasonable

A verification of the data base, again using canonical analysis, was undertaken with the univariate responders excluded. This procedure lowers the erroneously higher canonical statistics which tended to "cloud" the analysis using all respondents. Clearly, the redundancy between the variable sets is absolute for the univariate responders, so their exclusion should be expected to reduce the canonical correlations in the variable sets. The revised canonical correlation shows similarly high relation between the first pair of factors (.5528); the second pair (.4896); and the third pair (.4060). Overall these correlations are lower than previously measured, as expected.

#### Defining the Segments

It was not possible to segment the elderly market in this study on the usual a priori basis. We did not have knowledge about usage or non-usage of a specific product or product category or the criteria usually employed in designating heavy, medium or light users. Rather we were seeking a natural and more catholic segmentation, based on self concepts of buyer behavior. Ours is mainly a descriptive, rather than predictive, study. With no predefined segments, a clustering algorithm, Ward's method using an euclidean metric (Everett, 1974, p. 15) was employed. This clustering routine developed six segments based on the self-concept buyer behavior set. The criterion for using the six cluster solution was to strike a balance between minimizing the increase in variance resulting from combining respondents into clusters or clustering clusters to reduce their number (Anderberg, 1973) and the small cell sizes resulting from increasing the number of clusters produced.

Our analysis assumed an interval scale in measuring the buyer behavior characteristics and our initial description of each of the six clusters used the highest and lowest mean response for each characteristic. However, we assumed a null hypothesis that the mean response for each characteristic would be equal across clusters. Thus, a high level of significance of the F-test statistic would be grounds for rejecting the null hypothesis and accepting the alternative hypothesis that a significant difference does exist between the mean responses across groups. Therefore, we adopted a decision rule that any characteristic having an F-statistic significance greater than .10 be stricken from our analysis. This procedure eliminated the self-concept of impulsive buying (when in store, I often buy an item on the spur of a moment) with a significance of .1020. It was also decided to eliminate the final characteristic (all products that pollute the environment should be banned) since this was not a self-concept of buyer behavior, but rather an attitude toward a social problem.

The analysis produced behavioral descriptions of each segment that are relative, not absolute. The overall profile of these elderly consumers (Figure 2) should be used as a guide or norm. The cluster descriptions (Figs. 3-4) are relative to that overall elderly profile.

For the sake of identification we have collapsed those descriptions into the following more operational identifications:

#### Segment

1. Saver/planner
2. Brand Loyalist

3. Information Seeker
4. Economy Shopper
5. Laggard
6. Conspicuous Consumer

#### Segment Definition

The next step in our analysis was to broaden the definition of each of these segments by trying to describe each in terms of their corresponding demographic and psychographic characteristics:

#### Demographic Definitions

The TGI demographic variables are mostly categorical and the analysis applied to them was a chi-square test. We used a simple cross-tabulation of each demographic characteristic for each of the six behavioral clusters. The chi-square null hypothesis in this case is that the demographic variable is independent of cluster membership. Thus, the number of respondents in each cell is a function only of the marginal distributions of the variables, i.e., the demographic variable and the cluster membership variable. A significant chi-square statistic would reject the null hypothesis and lead to acceptance of the alternative hypothesis that a relationship does exist between the demographic variable and cluster membership. Once again our decision-rule was a significance figure of .10 or below to reject the null hypothesis. The data produced significance levels for chi-square considerably above .10. Thus, we cannot further describe the behavioral segments using demography at this time. However, because of the small cell sizes in the cross-tabulation we do not reject the possibility of demographic differences among clusters.

For those demographic variables which approximate an interval scale, or at least are of a higher than nominal scale of measurement, we ran Univariate one-way ANOVA's to detect significant mean response differences between clusters. Although none of these demographics showed significant differences between means, one of the variables showed some tendency toward discrimination. Education level (F-statistic significance of .175) was highest for group 3 (6.83) and lowest for group 2 (5.6) indicating that information seekers tend to have completed more formal education than brand loyalist consumers. This variable was coded as follows:

- 3 - some grammar school
- 4 - completed grammar school
- 5 - some high school
- 6 - completed high school
- 7 - some college
- 8 - completed college
- 9 - some graduate school
- 10 - completed graduate school

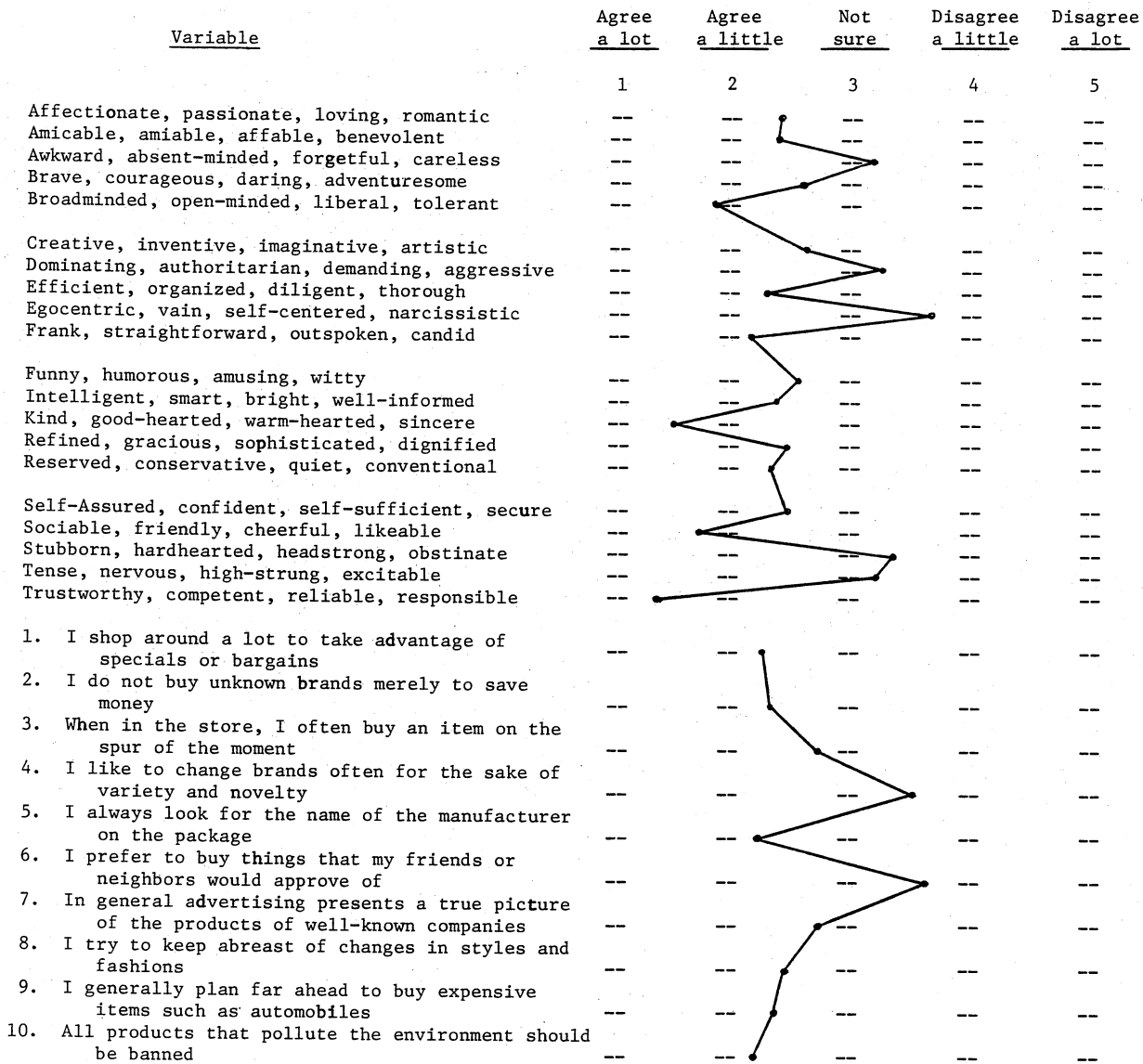
#### Psychographic Descriptions

In the case of the psychographic dimensions we again assumed an interval scale and used the highest and lowest mean response in our descriptive process. Consistently we used the decision rule of .10 or less for the F statistic significance to employ the psychographic characteristic in the analysis. This procedure eliminated eleven of the twenty self-concepts of psychographics from further descriptive use. The resultant descriptions for each buying style segment are detailed in Figure 4.



FIGURE 2

Overall Elderly Buyer Profile<sup>2</sup>



<sup>2</sup> Note; this profile represents the mean (average) response to each item across all elderly consumers in the sample excluding univariate responders. Use these mean responses as a guide when examining the definition and description of the behavioral segments. The behavioral segments are defined and described on a basis which is relative to this overall elderly buyer profile.

FIGURE 3  
Buying Style Cluster Definitions

<u>Cluster</u>	<u>Definition</u>
1	Buy unknown brands to save money Plan ahead to buy expensive items
2	Do not buy unknown brands Do not change brands for sake of variety Do not buy for approval of friends Brand Loyal
3	Not economy shopper Persuasive Style conscious
4	Economy shopper Not brand loyal
5	Not persuasive or style conscious Not a planner
6	Change brands Buy things friends approve of

FIGURE 4  
Psychographic Description of Elderly  
Buyer Clusters

<u>Cluster</u>	<u>Description</u>
1	Frank, candid Self-assured, confident, etc.
2	Brave, daring, courageous Reserved, conventional Not self-assured, not secure Not stubborn, not headstrong
3	Kind, sincere, good hearted
4	Not brave Not dominant Not egotistical Not frank Witty
5	Not witty Not kind Not reserved
6	Dominant Egotistical Stubborn

Figure 5 shows the matching of the buying style definitions of each cluster or segment to those psychographic descriptions and the relative size of each of these elderly sub-segments

#### Conclusions

The analysis identified six buying style segments of the elderly market and relates psychographic characteristics to each segment. The largest of these are identified as conspicuous consumers who self evaluate themselves as relatively more stubborn, egotistical and dominating than their consuming peers. The second largest segment are the saver/planners who tend to buy unknown brands and are self-described as more candid and confident. The fact that these two segments con-

FIGURE 5  
Buying Style Segments -- Psychographic Descriptions

<u>Segment</u>	<u>Psychographic Description</u>	<u>Share of Elderly Market (Percentage)</u>
Saver/Planner (buys unknown brands)	Frank, candid self-assured confident	25.1
Brand Loyalist (Does not buy for approval of friends)	Brave, courageous reserved, conventional Insecure not stubborn	8.4
Information Seeker (persuasive)	Kind, sincere	10.1
Economy Shopper (not brand loyal)	Not brave not dominating not egocentric not frank, candid funny, witty	10.6
Laggard (not persuasive)	Not witty not kind not reserivitive, liberal	11.2
Conspicuous Consumer	Stubborn Egotistical Dominating	34.6

stitute almost 60 per cent of the elderly consumers and that only 8.4 per cent are brand loyalists is significant to the marketing strategist.

However there are methodological cautions the authors give to that strategist in evaluating this report.

First, the clustering used in determining the elderly sub-segments used an algorithm which limits the shape of the resultant clusters imposing the same shape on all (Everitt, 1974, pp. 46-48).

Second, we recognize that the canonical analysis may capitalize on factors which are not necessarily the same factors which would be extracted in a principal components analysis (Cooley and Lohnes, 1971, p. 171).

Third, there is no consideration of a possible halo effect in the response (Wilkie, et al, 1974). We also recognize there may not be independence in responses, although the factor analytic routines used by ARMB in constructing the TGI data base were designed to offset this possibility.

In summary, this study did find the existence of natural segments of the elderly market defined by buying style characteristics and it did fit psychographic descriptions to those natural segments that are sensible and operational.

## References

- Alpert, Mark I., and Robert A. Peterson, "On Interpretation of Canonical Analysis," Journal of Marketing Research, May, 1972, pp. 187-92.
- Anderberg, Michael R., Cluster Analysis for Applications, (New York: Academic Press, 1973).
- Bass, Frank M., and W. L. Wilkie, "A Comparative Analysis of Attitudinal Predictions of Brand Preference," Journal of Marketing Research, October, 1973, pp. 262-69.
- Cooley, William W., and Paul R. Lohnes, Multivariate Data Analysis, (New York: John Wiley and Sons, Inc., 1971).
- Cox, Keith K., and Ben M. Enis, The Marketing Research Process, (Pacific Palisades: Goodyear Publishing Co. Inc., 1972).
- Engels, Franklin M., Information and Coding, (Scranton, Pennsylvania: International Textbook Company, 1971).
- Everitt, Brian, Cluster Analysis, (New York: John Wiley and Sons, Inc., 1974).
- Frank, Ronald E., and William F. Massey and Yoram Wind, Market Segmentation, (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1972).
- Ginter, James L., "The Effects of Normalization on the Multivariate Model," Advances in Consumer Research, Volume 1, Scott Ward and Peter Wright, eds., (Urbana, Illinois: Association for Consumer Research, 1974), pp. 302-309.
- Haley, Russel I., "The Implications of Market Segmentation," Conference Board Record, March, 1969.
- Kotler, Phillip, Marketing Management: Analysis, Planning and Control, 2nd ed. (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1972).
- Landon, E. Laird, "Self Concept, Ideal Self Concept, and Consumer Purchase Intentions," Journal of Consumer Research, September, 1974, pp. 44-51.
- Smith, Wendell R., "Product Differentiation and Market Segmentation as Alternative Marketing Strategies," Journal of Marketing, July, 1956, pp. 3-8.
- Wells, William D., ed., Life Style and Psychographics, (Chicago: American Marketing Association, 1974).
- Wilkie, William L., John McCann, and David J. Reibstein, "Halo Effects in Brand Belief Measurement: Implications for Attitude Model Development," Advances in Consumer Research, Vol. 1, Scott Ward and Peter Wright, eds. (Urbana, Illinois: Association for Consumer Research, 1974), pp. 280-290.

CONCEPTUAL AND OPERATIONAL ISSUES

IN THE EXTENDED FISHBEIN MODEL

Richard J. Lutz, University of California, Los Angeles

Abstract

Fishbein's (1967a) model for the prediction and explanation of specific behaviors has attracted considerable attention among consumer researchers. The purpose of the present paper is to summarize some conceptual and operational issues surrounding the constructs in the model. In particular, attitude-toward-the-behavior and the relatively recent subjective norm construct are examined, and possible reconceptualizations are offered. Measurement problems with respect to the cognitive components in the model are discussed, and suggestions are made for the improved testing of model assumptions. Model testing and refinement are advocated, and appropriate research strategies, allowing causal analyses, are outlined.

Fishbein's (1967a, 1975) modification of the theory of propositional control (Dulany, 1968) has emerged over the past few years as one of the more popular theories in consumer research. This popularity is at least in part due to the fact that the theory is a promising one for consumer researchers interested primarily in the explanation of consumer behavior and, at the same time, provides a framework for devising behavior change strategies, a feature of interest to managers and policy makers.

In their recent review of research generated by this theory in the consumer domain, Ryan and Bonfield (1975) have provided an excellent summary of empirical findings, as well as raising some important research issues. The purpose of the present investigation is to discuss some additional conceptual and operational issues with respect to the Fishbein model. To accomplish this task, the most recent form of the model will be presented briefly, followed by detailed consideration of issues surrounding the various theoretical constructs in the model.

The Extended Fishbein Model

The Extended Fishbein Model (so designated because it is viewed as an extension of Fishbein's (1963) earlier model of attitude) incorporates attitudinal and social influences in attempting to explain the formation of behavioral intention, which is seen as the immediate predecessor of overt behavior. While nomenclature and operational specification of the constructs in the model have undergone several changes since the original formulation in 1967, Fishbein and Ajzen's (1975) recent book provides an update which will serve as the basis for the present discussion. Essentially, the model rests on three equations:

$$B \rightsquigarrow I = (A_B)w_1 + (SN)w_2 \quad (1)$$

$$A_B = \sum_{i=1}^n b_i e_i \quad (2)$$

$$SN = \sum_{j=1}^m b_j m_j \quad (3)$$

where B is the behavior under study; I is the individual's intention to perform the behavior;  $A_B$  is his attitude toward performing the behavior; SN is the subjective norm with respect to the behavior;  $w_1$  and  $w_2$  are empirically derived weights;  $b_i$  is the belief that performing the behavior leads to some consequence i;  $e_i$  is the individual's evaluation of consequence i; n is the number of salient beliefs;  $b_j$  is the normative belief regarding referent j's expectations as to whether the individual should or should not perform the behavior;  $m_j$  is the individual's motivation to comply with referent j; and m is the number of salient referents (Fishbein and Ajzen, 1975, pp. 301-2).

Equation 1 is a theoretical statement of the determinants of intention, which is seen as more or less equivalent to behavior, depending upon several factors (e.g., the timing and specificity of measurement); Equation 2 specifies cognitive determinants of attitude. Equations 1 and 2 are familiar to consumer researchers; however, Equation 3, which deals with the determinants of the subjective norm, is a more recent addition to the model. Originally there was no SN construct in the model, and the summative compound shown on the right side of Equation 3 was used directly in Equation 1 to represent normative influences. Thus SN has the same logical status in the model as does  $A_B$  on the attitudinal side--both are summary measures of potential sources of influence (i.e., attitudinal and normative) on the formation of intentions.

Attention is now focused on issues surrounding each of the equations in the model, beginning with the attitudinal component (Eqn.2), followed by the normative component (Eqn. 3) and the intention equation. In each instance, issues have been separated into conceptual and operational categories to facilitate discussion.

The Attitudinal Component

Two major conceptual issues arise with respect to the attitudinal component, and both are concerned with the nature of the cognitive determination of  $A_B$ . First, the status of the  $b_i$  element as a probability dimension is considered; and second, the question of cognitive adding vs. cognitive averaging is discussed.

Status of the belief element. In developing his original  $A$  (attitude-toward-the-object) model, Fishbein (1967b) relied on the process of mediated generalization in behavioristic learning theory. Specifically, he proposed that the evaluative responses associated with certain salient characteristics of an object are transferred, or generalized, to the object itself, thus forming the individual's attitude toward the object. Further, the degree to which the evaluative response to any attribute was transferred to the object was assumed to be a function of the probability that the object possessed the attribute. Hence, a multiplicative summation identical to Equation 2 was the theoretical basis for  $A_O$ .

The attitude-toward-the-behavior ( $A_B$ ) construct has never undergone independent theoretical development;

rather, Fishbein (1967) posited that the same processes underlying the formation of  $A_B$  should also be the foundation of  $A_B$ . However, the types of beliefs ( $b_i$ ) which tend to be salient for  $A_B$  are no longer the descriptive beliefs salient for  $A_B$ , but rather are beliefs about the outcomes or consequences of engaging in the behavior. Thus, the "attributes" associated with the behavior are a set of outcomes, or valued states, which the individual wishes to either avoid or attain. The result of this rather subtle shift from descriptive beliefs to beliefs about behavioral consequences is that the attitude in question ( $A_B$ ) is one which has much more functional significance for the individual. That is, an individual's attitude toward engaging in a behavior will be favorable if the behavior leads to desirable outcomes or blocks undesirable outcomes for the individual. The key question then becomes whether the individual's beliefs about behavioral consequences take the form of subjective probability statements. There is some theoretical disagreement on this point.

Working within the functionalist tradition, Rosenberg (1956) developed his version of the "expectancy-value" model based on the assumption that the expectancy or belief component represented the degree to which the attitude object led to the blocking or attainment of certain values. Thus, rather than being concerned with the probability of association between an object and an outcome, Rosenberg dealt with the degree of association or dissociation between object and outcome.

When one analyzes the "attitude objects" studied by Rosenberg and the other functionalists, an interesting trend emerges: virtually every "object" studied was actually a behavior (see Table 1). While the behaviors under investigation were not generally those that the individual himself would have performed (e.g., most of the subjects in those studies would have had little power to remove Negro segregation), nevertheless those actions would have had the effect of blocking or facilitating the attainment of certain goals by the individual. Hence, the bipolar construct of "blocking/attainment" or "association/dissociation" seems to be a reasonable one for representing beliefs about behavioral consequences.

TABLE 1

REPRESENTATIVE "ATTITUDE OBJECTS" USED IN RESEARCH BY FUNCTIONAL THEORISTS

Reference	"Attitude Object"
Rosenberg (1956)	"allowing members of the Communist part to address the public" (p. 368)
Woodruff and DiVesta (1948)	"the proposed abolishment of fraternities and sororities from American colleges" (p. 649)
Carlson (1956)	"the removal of Negro segregation" (p. 130)
DiVesta and Merwin (1960)	"teaching as a career" (p. 81)
Peak (1960)	Eleven educational policies (e.g., "giving unannounced quizzes") (p. 464)
Axelrod (1963)	Using eleven products (e.g., "drinking a Daiquiri") (p. 22)
Rosenberg (1960)	"Negroes moving into White neighborhoods" (p. 45)

Fishbein and Ajzen (1975) have pointed out that the association/dissociation construct can be represented as two probabilistic belief dimensions. This is certainly true logically and testable empirically. But which approach comes closer to representing the form of beliefs individuals actually hold in cognitive structure? Intuitively, it seems more plausible that association/dissociation would be operative in beliefs with respect to behaviors. Recent empirical evidence tends to support that notion (Bettman, Capon and Lutz, 1975a,c).

As conceptualized by Fishbein, the  $b_i$  construct is a subjective probability dimension; it is typically measured on a 7-point scale ranging from "unlikely" to "likely". Fishbein and Ajzen (1975) advocate scale values for this measure of -3 to +3, thus treating it as a bipolar construct, which is hardly a probability statement. Interestingly, Bettman, et al. (1975a) found that subjects tend to respond to the  $b_i$  measure in a bipolar fashion, thus providing justification for the use of the bipolar coding scheme in data analysis. Whether this response bipolarity is a function of respondents' "true" cognitive structures or is an artifact of the measurement procedure is uncertain. It is clear, however, that in its present form the  $b_i$  construct does not behave as a probability dimension.

Ahtola (1975) has suggested a model for partitioning the  $b_i$  dimension into belief content and belief strength, thereby taking into account both the blocking/attaining feature and the probability aspect of  $b_i$ . This model, while having theoretical elegance, is quite cumbersome operationally. Further, it seems unlikely that individuals' cognitive structures are as differentiated as the Ahtola model implies. Instead, it seems more plausible that for the vast majority of beliefs, particularly those which are characterized by monotonic affect functions (Ahtola, 1975), belief content is the salient construct in cognitive structure. That is, the individual holds a belief regarding the degree to which the behavior in question leads to or blocks a particular goal. It is this belief, a bipolar construct, which should comprise the  $b_i$  component in the Extended Fishbein Model. The probability dimension does not adequately capture the idea of association/dissociation, and consideration of both probability and association/dissociation overstates the information processing capabilities of most individuals.

The above discussion argues against none of the empirical research on the Fishbein model, all of which has treated  $b_i$  as a bipolar construct, though in theory it is a (unipolar) probability. The purpose of the discussion is simply to point out that the model has worked well, as operationalized, which happens to conflict with the model as conceptualized. What is argued, therefore, is that the weight of empirical evidence suggests a change in the conceptualization of the  $b_i$  component to bring it more into line with empirical evidence. Once the  $b_i$  dimension is accepted as being conceptually bipolar, then appropriate measurement procedures can be adopted to more accurately reflect that dimension. Bettman, et al. (1975c) were able to increase the strength of the relationship between  $A_B$  and  $E_{b_i}e_i$  by substituting a "high-low" scale for the commonly used "likely-unlikely" scale. The assumption was that subjects would be better able to respond to a scale which more closely approximated their own belief structures. Thus, "high-low" was seen as being more inherently bipolar than "likely-unlikely" and if subjects' beliefs tended to be bipolar, they should have been able to respond with less difficulty to the former scale. While the results were not conclusive, there was a good indication that the "high-low" scale performed better. Obviously, there may be other scale endpoints which would be superior to

"high-low". Nevertheless, there is ample justification for questioning the status of the  $b_i$  construct as a probability dimension.

Further research should be directed at uncovering the nature of the  $b_i$  component. One approach would be the analysis of free response data generated by respondents to determine how they seemingly organize belief statements. The use of multitrait-multimethod procedures would also appear to be a valuable tool in this task.

Adding vs. averaging. Fishbein (1965), in discussing his  $A_B$  model, states: "...the assumption of 'summation' is not explicit in behavior theory. Thus, for example, a postulation of 'averaging' would also have been consistent with behavior theory" (p. 117). Much research has been devoted in the psychological literature to determining whether beliefs add or average to form overall attitude. Probably the major proponent of an averaging formulation is Anderson (1971), who has demonstrated averaging effects within his integration theory approach to cognitive processes. Fishbein and Ajzen (1975), however, dismiss most of the research on the adding/averaging issue on methodological grounds and prefer to stand by the summation assumption in Equation 2.

Bettman, *et al.* (1975b) found evidence that most of their subjects tended to average rather than add across beliefs, within the context of Fishbein's  $A_B$  model. Thus, a serious question is raised regarding the appropriateness of the original summation assumption. The issue is unimportant when static tests of the model are undertaken, as the average is simply a linear transformation of the sum and is therefore perfectly correlated with it. However, in attitude change situations, the distinction is a crucial one, for the adding and averaging formulations often make opposite predictions. For instance, under a summation model, a piece of moderately favorable information added to a highly favorable initial attitude would be expected to increase the favorability of the attitude. However, an averaging assumption would predict a decrease in the favorability of the attitude. Obviously, this is an important concern for policy makers using the model for generating attitude change strategies.

Future research should use an attitude change paradigm to determine whether averaging or adding is the dominant cognitive process underlying  $A_B$ . Undoubtedly, individual differences and situational influences may have impact on this process, and to the extent that they do, there can be no global generalization. In light of recent evidence, it seems inappropriate to continue with the assumption that Equation 2 is correct as it stands.

Operational Issue. The major operational issue confronting the Extended Fishbein Model, as well as all other expectancy-value formulations, is the problem with the scale properties of the expectancy ( $b_i$ ) and value ( $e_i$ ) terms. Schmidt (1973, 1975) has demonstrated that in the absence of ratio measure for both  $b_i$  and  $e_i$ , the correlation between  $A_B$  and  $\Sigma b_i e_i$  can be artificially manipulated to virtually any value by simple transformations of  $b_i$  and  $e_i$ . These transformations would not be possible on ratio data, but are perfectly legitimate on interval data. Therefore, from a mathematical perspective,  $\Sigma b_i e_i$  is a meaningless computation unless ratio data are obtained.

The problems inherent to obtaining ratio scale data are well known; it is safe to say that none of the hundreds of studies conducted on the various expectancy-value models in the literature have used such data. This is a

serious concern, for it means that standard correlation approaches to model validation are essentially worthless in the absence of other kinds of data.

Table 2 shows the correlations calculated between  $A_B$  and  $\Sigma b_i e_i$  for different combinations of coding assumptions. Data for these correlations came from 491 housewives responding to a questionnaire about a hypothetical laundry detergent.<sup>1</sup>

TABLE 2  
PRODUCT-MOMENT CORRELATIONS BETWEEN  
 $A_B$  AND  $\Sigma b_i e_i$  WITH VARYING CODING ASSUMPTIONS

Coding for $e_i$	Coding for $b_i$				
	-3/+3	-2/+4	-1/+5	0/+6	+1/+7
-3/+3	.445	.451	.445	.433	.418
-2/+4	.455	<u>.457</u>	.451	.438	.423
-1/+5	.445	.447	.442	.432	.419
0/+6	.425	.428	.425	.418	.408
+1/+7	.403	.407	.406	.401	.394

The upper lefthand entry in the matrix is the correlation under the typical coding assumptions of -3/+3 for both the  $b_i$  and  $e_i$  components. The underlined entry in Table 2 is the maximum correlation attained under any of the coding schemes employed. This corresponds to a coding scheme wherein subjects' responses were coded from -2 to +4 for both  $b_i$  and  $e_i$ . The fact that the correlation is slightly higher under the latter approach suggests that subjects were displacing the neutral point on both scales toward the negative endpoint.

An identical conclusion was reached by Bettman, *et al.* (1975a) using an ANOVA framework for investigating the model: In their data, an expected symmetric interaction pattern was shifted toward the negative ends of the scales, indicating that a -3/+3 coding scheme tends to underestimate the favorability of the  $b_i e_i$  contribution to overall attitude.

In another context, Bentler (1969) has shown semantic space to be approximately bipolar. Since the good-bad and likely-unlikely scales used by Fishbein to measure  $e_i$  and  $b_i$  are derived from the work of Osgood, *et al.* on the semantic differential technique, it seems plausible that the  $b_i$  and  $e_i$  scales should also be approximately bipolar. This is in agreement with the empirical results discussed above.

Thus, while there is no direct evidence pertaining to the scale properties of the  $b_i$  and  $e_i$  measures, it appears that they are at least fairly adequately represented by the -3/+3 coding scheme advocated by Fishbein and Ajzen (1975). Nevertheless, there is an indication that a slightly asymmetric coding procedure would be more appropriate.

The only strictly correct operational procedure for testing the relationship between  $A_B$  and  $\Sigma b_i e_i$ , as currently conceptualized by Fishbein and Ajzen (1975), would be to obtain ratio measures of  $b_i$  and  $e_i$ . For the prob-

<sup>1</sup>For a more detailed description of these data, the reader is referred to Lutz (1975a).

ability component, this would be a relatively easy task, as subjects could be asked to provide subjective probability estimates on a "chances out of 100" scale. (Incidentally, data of this type would be necessary to refute the arguments against the use of subjective probabilities which were presented in the preceding section of this paper.)

Measuring  $e_i$  on a ratio scale is a more imposing task. As yet, no completely satisfactory procedure exists. Thurstone and Jones (1957) have presented a model for determining the "rational origin" for subjective values, but the method is quite cumbersome and offers no guarantee that an origin can be located in every instance.

Adequate operationalization of the  $b_i$  and  $e_i$  constructs in Equation 2 as subjective values with a natural origin is the only procedure by which the Fishbein model can be tested correlationally. ANOVA procedures allow tests without making scale assumptions regarding the  $b_i$  and  $e_i$  components, but are fraught with problems of external validity (Bettman, Capon and Lutz, 1975d). It appears that further tests of the validity of Equation 2 in the model must rely on multimethod approaches where correlational and ANOVA procedures are combined, wherever possible within attitude change contexts.

#### The Normative Component

Conceptual Issues. The normative component of the Extended Fishbein Model has always been its weak link. Only recently has the subjective norm (SN) construct been formulated, and empirical evidence in support of it is sparse. Similarly, the theoretical antecedents of SN are poorly developed conceptually. Therefore, both sides of Equation 3 raise major conceptual issues: both the criterion variable SN and its antecedents,  $b_j$  and  $m_j$ , need fuller development.

Subjective norm. The normative component has suffered from the lack of an overall criterion measure (similar to  $A_B$  for the attitudinal component) since the genesis of the model. Absence of this construct has prevented serious validation work on the determinants of normative pressure, i.e., normative beliefs ( $b_j$ ) and motivation to comply ( $m_j$ ).

In their recent book, Fishbein and Ajzen (1975) report the use of the following scale to measure SN:

Most people who are important to me think  
I should \_\_\_\_\_ I should not  
perform behavior X.

Correlations ranging from .625 to .910 with  $\Sigma b_j m_j$  were found; yet, the above statement does not seem to capture the essence of motivation to comply. Rather, the above scale appears to measure a generalized normative belief, i.e., what others think the person should do. If the phrase, "who are important to me," is meant to somehow pick up  $m_j$ , then it seems inadequate. There is little chance that "negative" referents, for whom  $m_j$  would be small or even in the opposite direction, would be evoked by this phrase.

Ryan (1975) has recently proposed a social compliance (SC) construct similar in purpose to SN and has used Kelman's (1958) notions as a conceptual basis. Specifically, a scale was constructed to measure only the compliance process in Kelman's typology, the assumption being that the processes of identification and internalization should be captured by the attitudinal component. Initial empirical work showed a statistically

significant correlation SC and  $\Sigma b_j m_j$ ; however, SC also correlated highly with  $\Sigma b_j e_j$  and  $A_B$ , suggesting the need for further work on this construct.

The essential conceptual content which must be captured by the SN construct is the general social pressure on the person to perform the behavior. Thus, both normative beliefs and the motivation to comply with these beliefs should be represented in this single unidimensional construct. One way to conceptualize SN is to treat it in the same manner as  $A_B$ .  $A_B$  is an overall affective dimension, where the intensity of affect is seen as deriving from a number of affective ( $e_i$ ) dimensions (Fishbein, 1963). The degree to which each of the  $e_i$  dimensions contributes to  $A_B$  is a function of belief strength ( $b_i$ ).

To treat SN analogously, the SN construct would be viewed as a global motivation to comply with other people in deciding whether or not to perform the behavior in question. The source of this motivation to comply would be the individual  $m_j$  elements which represent motivation to comply with certain referents, in general. To the extent that these referents are perceived as having beliefs about what the individual should do with respect to the behavior in question, then the various  $m_j$  elements will combine to form a general motivation to comply with others in that situation. Under this approach, SN might be measured on a scale such as:

With respect to performing behavior X,  
I very much \_\_\_\_\_ I very much  
want to do \_\_\_\_\_ want not to do  
what other people think I should do.

To the extent that intention (I) is under normative control in a given situation, then SN should be highly correlated with I. Note that under this approach it would not be necessary to represent directly in SN the direction of normative beliefs; all that is important is the degree to which the individual is motivated to comply with these beliefs. The regression weight ( $w_2$ ) attached to the normative component would reflect the direction of normative beliefs, in general. This is exactly the same approach as is used in measuring  $A_B$ : affect is the only important feature, not the beliefs from which that affect derives. In order to use the normative component for diagnosis, individual  $b_j$  elements would have to be analyzed, just as is presently done with  $b_i$  elements underlying  $A_B$ .

Thus the determinants of intention are viewed as two response tendencies of the individual. One response tendency is attitudinal and reflects the individual's own personal desire to engage in the behavior. The other response tendency is normative and summarizes the person's motivation to go along with others in deciding whether or not to perform the behavior. These two response tendencies may be competing or complementary and should in general behave in accordance with the propositions outlined by Fishbein (1967a, 1975).

It is not recommended that a single scale such as the one shown above be used to measure SN. Rather, a series of items, in either semantic differential or Likert format, should be employed to achieve a reasonable degree of reliability. Again, Ryan's (1975) measure of SC is a useful first step in this direction.

Determinants of subjective norm. Perhaps due to the absence of the SN construct until recently, there has been considerable ambiguity surrounding the exact nature of the normative belief ( $b_j$ ) and motivation to comply

( $m_j$ ) elements underlying the normative component. In its most recent form (Fishbein and Ajzen, 1975),  $b_j$  is treated in a fashion similar to  $b_i$  for the attitudinal component. Thus a normative belief ( $b_j$ ) is seen as a probabilistic statement that a particular referent thinks the individual should or should not perform the behavior in question. Motivation to comply ( $m_j$ ) with the normative belief is defined as the general tendency of the individual to go along with the referent's expectations, although there is some sentiment for adopting a more situation specific conceptualization of  $m_j$  (Fishbein and Ajzen, 1975, p. 306).

A useful approach to more clearly delineating the conceptual properties of  $b_j$  and  $m_j$  might be to consider, separate from  $A_B$ , the social consequences of performing the behavior in question. Thus the individual may be asked to respond to  $b_j$  statements like:

If I purchase Brand X, my family will hate me

likely \_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:unlikely

If I purchase Brand X, my family will cut off my allowance

likely \_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:unlikely<sup>2</sup>

Note that this procedure would permit a multidimensional representation of any particular referent's relationship to the individual, thus allowing possible incorporation of Kelman's (1958) typology of social influence and French and Raven's (1959) bases of social power.

Corresponding to the above  $b_j$  dimensions would be  $m_j$  dimensions more closely akin to the  $e_i$  elements underlying  $A_B$ . Thus,

For my family to have me is

good \_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:bad

For my family to cut off my allowance is

good \_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:bad

represent evaluations of the social consequences of engaging in the behavior. The impact of these changes in  $b_j$  and  $m_j$  on the SN construct would be to convert it into an attitude toward engaging in the behavior due to social demands (rather than personal evaluations). This may be thought of as similar to Rokeach's (1968) notion of attitude toward the situation. Thus, intention would be determined by the interaction of two attitudes-- $A_B$ , which is based on personal beliefs about the consequences of engaging in the behavior, and  $SA_B$  (social attitude toward the behavior), which is an attitude toward engaging in the behavior based exclusively on social outcomes.

Whether these two attitudes can be operationalized such that they are independent of each other is an empirical question. The important point of the above discussion is that the normative component can be conceptualized using the same basic elements (i.e., beliefs and evaluations) as are used in deriving  $A_B$ . This approach is at least worthy of some consideration, given the rela-

<sup>2</sup>While it will not be discussed again here, arguments against the use of probability dimensions in conceptualizing the  $b_i$  component of  $A_B$  apply equally to the use of probability to represent  $b_j$ . The likely-unlikely scales are used here to show the similarity between the construct being proposed here and the  $b_i$  component as currently operationalized.

tive success of the attitudinal component in past research and the general lack of success with respect to the normative component.

Operational Issue. The chief operational issue with respect to the normative component of the Extended Fishbein Model is similar to the problem encountered in scaling  $b_i$  and  $e_i$  elements. Both  $b_j$  and  $m_j$  should be ratio scales to allow multiplication of the terms as specified in Equation 3. In practice, both elements underlying the normative component have typically been measured on 7 point scales, coded -3/+3. Schmidt's (1973, 1975) analysis applies to this practice, as well as the coding procedures for the attitudinal component which were discussed earlier. Correlations can be manipulated drastically by simply changing the coding assumptions. Table 3 shows a wide range in correlation between  $E_{b_j m_j}$  and intention, based on the same data set as the attitudinal correlations reported earlier in this paper. Intention was used as the dependent measure, as this research was conducted prior to the formulation of the SN construct.

TABLE 3

PRODUCT-MOMENT CORRELATIONS BETWEEN INTENTION AND  $E_{b_j m_j}$  WITH VARYING CODING ASSUMPTIONS

Coding of $m_j$	Coding of $b_j$				
	-3/+3	-2/+4	-1/+5	0/+6	+1/+7
-3/+3	.107	.144	.156	.156	.152
-2/+4	.362	.363	.334	.302	.274
-1/+5	.487	.471	.435	.395	.359
0/+6	.534	.517	.485	.450	.416
+1/+7	.551	.536	.511	.482	.452

In this case,  $b_j$  was measured on scales like the following:

My family would \_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_ would not

expect me to try Brand M detergent.

Motivation to comply ( $m_j$ ) was a situation specific measure, as follows:

With respect to my choice of laundry detergent

I want to \_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_ do not want to

do what my family expects me to do.

As can be seen in Table 3, the assumption of -3/+3 coding for both elements is incorrect (if correlation with intention can be accepted as a criterion). While  $b_j$  appeared to operate in a bipolar manner,  $m_j$  was clearly treated as a unipolar construct by the respondents. Hence, a 1 to 7 coding scheme for  $m_j$  increases the correlation between  $E_{b_j m_j}$  and I to its maximum (lower lefthand entry in Table 3), while the standard coding procedure results in the minimum correlation between the two variables (upper lefthand entry in the table).

Obviously, much more work must be devoted to obtaining satisfactory measures of  $b_j$  and  $m_j$ . The likelihood of obtaining ratio scales seems quite small; however, the use of ANOVA procedures similar to those used for the attitudinal component (e.g., Bettman, et al., 1975a) should provide useful insights.



## Determinants of Intention

**Conceptual Issues.** Conceptual issues with respect to the overall intention equation (Eqn. 1) concern 1) the status of intention (I) as a variable mediating the effects of all other cognitive constructs on behavior (B), and 2) the psychological meaning of the weights ( $w_1$  and  $w_2$ ).

**Status of intention.** Intention is a familiar construct to consumer researchers. It has been used at an aggregate level since the early 1950's when Katona and his associates at the University of Michigan's Survey Research Center began incorporating an intention variable in their economic forecasts. Recent consumer theories (e.g., Howard and Sheth, 1969) have postulated an intention variable as the immediate precursor of purchase behavior and have treated it as an explanatory construct for individual purchase behavior.

In a similar vein, Fishbein (1967) has elevated intention from the status of being one of the components of attitude (i.e., conation) to a construct in its own right. Equation 1 above clearly specifies the role of intention as a mediating, summary cognitive construct in which attitudinal and normative factors are combined to reach a decision to behave in a certain manner. Thus, a causal flow of influence is postulated wherein I is the resultant of  $A_p$  and SN and, in turn, is the antecedent of B (Ryan, 1975). Substantiation of these proposed relationships is essential if the Extended Fishbein Model is to be used in the formulation of behavior change strategies.

Recent research on causal patterns within the Fishbein model (Lutz, 1975b; Ryan, 1975) has shown promising results for the hypothesized relationships. Neither of these studies measured overt behavior, however. Further research using change experiments and longitudinal designs should be undertaken, and, wherever possible, measures of overt behavior should be obtained.

**Psychological meaning of  $w_1$  and  $w_2$ .** The weights attached to the attitudinal and normative components in Equation 1 have traditionally been estimated through regression procedures. Fishbein and Ajzen (1975) state: "Ideally, the weights for the attitudinal and normative components would be available for each individual with respect to each behavior in a given situation. Since adequate estimates of this kind are not presently available, the practice has been to use multiple regression techniques..." (p. 303).

But what do the weights represent conceptually? If they represent the relative importance that the individual places on attitudinal vs. normative considerations, then surely such estimates could be obtained. Of course, such judgments would not be very reliable, as recent clinical judgment research has shown. Nevertheless, a conceptual explanation of the weights seems necessary before it can be concluded that independent estimates of their magnitudes are not available. The present regression procedure has several drawbacks, as shall be seen in the next section.

**Operational Issues.** Three areas of concern arise with respect to the operationalization of Equation 1: first, measurement of intention can be improved; second, there is a persistent problem of multicollinearity between the attitudinal and normative components; and third, more work needs to be done on the testing of the model at the individual level.

**Measurement of intention.** Juster (1966) reports that a purchase probability scale (an 11-point scale ranging from "0 chances out of 10" to "10 chances out of 10") was "markedly superior" (in the prediction of behavior) to traditional measures of purchase intention. Little attention has been paid to the measurement of I within the Fishbein model. Typically, multiple item semantic differential inventories are used to provide a single score; however, there has been no attempt made to relate this score to the concept of purchase probability. While Juster's (1966) results were based on aggregate level analyses, his conclusion about the usefulness of the probability scale suggests that a similar form of measurement may be useful for the explanation of individual purchase behavior.

**Multicollinearity of  $A_p$  and SN.**<sup>3</sup> It is well known that high correlations between predictor variables in a multiple regression cause the regression weights attached to those variables to become unstable. This is an unsatisfactory situation for the manager interested in basing strategy decisions on the relative magnitude of those weights. Yet in many cases, due to a halo effect present in the measurement procedure or perhaps a true relationship between the two components,  $A_p$  and SN have been highly correlated. Under such conditions, the Extended Fishbein Model is weakened for explanatory purposes. This points to the need, discussed in the previous section, for independent assessment of  $w_1$  and  $w_2$ .

**Individual level analysis.** Most tests of Equation 1 have been conducted via cross-sectional regressions. This procedure is not completely satisfactory, as there is no information provided regarding individuals'  $w_1$  and  $w_2$  parameters. In order to understand consumer behavior at the individual level, a separate set of weights for each person in the sample would be desirable. Recently, Wilson, et al. (1975) and Kakkar (1975) have reported the results of individual level analyses which support the Fishbein model. Wilson, et al. were able to find significant clusters of respondents with respect to the patterns of their regression weights, which suggests the possibility of segmentation analysis. However, individual level analysis may be a two-edged sword; Kakkar (1975) found heightened multicollinearity between  $A_p$  and  $\sum b_j m_j$  in his study. Wilson, et al. do not report the correlation between  $A_p$  and  $\sum b_j m_j$ .

In order to test the model at the individual level without the threat of multicollinearity or the problems with scale properties, Anderson's (1971) information integration framework can be employed. Similar to the studies reported by Bettman, Capon and Lutz (1975a,b,c,d), a factorial design incorporating  $b_i$ ,  $e_i$ ,  $b_j$  and  $m_j$  as four completely crossed factors and intention as the dependent variable would allow tests of several assumptions. First, the multiplicative relationship of  $b_j$  and  $m_j$  could be examined (the multiplicative relationship of  $b_i$  and  $e_i$  has already been established). Second, the assumption that the attitudinal and normative components are additive, as shown in Equation 1, could be tested. This would provide evidence as to whether the observed multicollinearity between  $A_p$  and SN is due to measurement effects or a true configural relationship between the two components.

## Conclusion

Although a relatively recent addition to the consumer

<sup>3</sup>Strictly speaking, most instances of multicollinearity have been between  $A_p$  and  $\sum b_j m_j$ , since SN has only recently been included in the model. However, Ryan (1975) observed multicollinearity between  $A_p$  and SC.

literature, the Extended Fishbein Model has shown considerable promise as both a managerial and research tool. Recent developments with respect to the normative component point to the need for further testing and refinement of the model.

As was true with early research on the multiattribute attitude model, much of the research on the Extended Fishbein Model has been within a static, cross-sectional paradigm. In order to test a theory as rich as the Fishbein theory, other approaches are essential. In particular, dynamic tests of the model are useful for investigating causal relationships. In a similar vein, longitudinal designs would help to identify the intricate feedback mechanisms between behavior and beliefs, behavior and attitude, etc. Continued emphasis should be placed on individual level analysis, and ANOVA procedures should be used to complement the more commonly used correlational designs.

#### References

- Ahtola, Olli T. "The Vector Model of Preferences: An Alternative to the Fishbein Model," Journal of Marketing Research, 12 (February 1975), 52-59.
- Anderson, Norman H. "Integration Theory and Attitude Change," Psychological Review, 78 (1971), 171-206.
- Axelrod, Joel N. "Induced Moods and Attitudes Toward Products," Journal of Advertising Research, 3 (June 1963), 19-24.
- Bentler, P.M. "Semantic Space is (Approximately) Bipolar," Journal of Psychology, 71 (1969), 33-40.
- Bettman, James R., Noel Capon, and Richard J. Lutz. "Cognitive Algebra in Multi-Attribute Attitude Models," Journal of Marketing Research, 12 (May 1975a), 151-64.
- \_\_\_\_\_. "Information Processing in Attitude Formation and Change," Communication Research, 2 (3) (July 1975b), 267-78.
- \_\_\_\_\_. "Multiattribute Measurement Models and Multiattribute Attitude Theory: A Test of Construct Validity," Journal of Consumer Research, 1 (March, 1975c), 1-15.
- \_\_\_\_\_. "A Multimethod Approach to Validating Multi-Attribute Models." In M.J. Schlinger (Ed.), Advances in Consumer Research, Volume 2. Chicago: The Association for Consumer Research, 1975d, 357-74.
- Carlson, Earl R. "Attitude Change Through Modification of Cognitive Structure," Journal of Abnormal and Social Psychology, 52 (1956), 256-61.
- Divesta, Francis J., and Jack C. Merwin. "The Effects of Need-Oriented Communications on Attitude Change," Journal of Abnormal and Social Psychology, 60 (1960), 80-5.
- Dulany, Don E. "Awareness, Rules, and Propositional Control: A Confrontation with S-R Behavior Theory," In Theodore R. Dixon and David L. Horton (Eds.), Verbal Behavior and General Behavior Theory. Englewood Cliffs, N.J.: Prentice-Hall, 1968, 340-87.
- Fishbein, Martin A. "An Investigation of the Relationships Between Beliefs About an Object and the Attitude Toward That Object," Human Relations, 16 (August 1963), 233-9.
- \_\_\_\_\_. "A Consideration of Beliefs, Attitudes, and Their Relationship." In Ivan D. Steiner and Martin Fishbein (Eds.), Current Studies in Social Psychology. New York: Holt, Rinehart, and Winston, 1965, 107-20.
- Fishbein, Martin A. "Attitude and the Prediction of Behavior." In Martin Fishbein (Ed.), Readings in Attitude Theory and Measurement. New York: Wiley, 1967a, 477-92.
- \_\_\_\_\_. "A Behavior Theory Approach to the Relations Between Beliefs About an Object and the Attitude Toward the Object," in Martin Fishbein (Ed.), Readings in Attitude Theory and Measurement. New York: Wiley, 1967b, 389-400.
- \_\_\_\_\_, and Icek Ajzen. Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Reading, Mass.: Addison-Wesley, 1975.
- French, J.R.P., Jr., and Bertram Raven. "The Bases of Social Power," in D. Cartwright (Ed.), Studies in Social Power. Ann Arbor: University of Michigan Press, 1959.
- Howard, John A., and Jagdish N. Sheth. The Theory of Buyer Behavior. New York: Wiley, 1969.
- Juster, F. Thomas. "Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design," Journal of the American Statistical Association, 61 (September 1966), 658-96.
- Kakkar, Pradeep. Situational Influence on Emotional Response, Cognitive Structure and Cognitive Process. Unpublished doctoral dissertation, University of California, Los Angeles, 1975.
- Kelman, Herbert C. "Processes of Opinion Change," Public Opinion Quarterly, 25 (Spring 1961), 57-78.
- Lutz, Richard J. "Changing Brand Attitudes Through Modification of Cognitive Structure," Journal of Consumer Research, 1 (March 1975a), 49-50.
- \_\_\_\_\_. "An Experimental Investigation of Causal Relations Among Cognitions, Affect and Behavioral Intention," Working Paper No. 34, Center for Marketing Studies, University of California, Los Angeles, 1975b.
- Osgood, Charles E., George J. Suci, and Percy H. Tannenbaum. The Measurement of Meaning. Urbana: University of Illinois Press, 1957.
- Peak, Helen. "The Effect of Aroused Motivation on Attitudes," Journal of Abnormal and Social Psychology, 61 (1960), 463-8.
- \_\_\_\_\_, and H. William Morrison. "The Acceptance of Information into Attitude Structure," Journal of Abnormal and Social Psychology, 57 (1958), 127-35.
- Rokeach, Milton. Beliefs, Attitudes and Values: A Theory of Organization and Change. San Francisco: Jossey-Bass, 1968.
- Rosenberg, Milton J. "Cognitive Structure and Attitudinal Effect," Journal of Abnormal and Social Psychology, 53 (1956), 367-72.
- \_\_\_\_\_. "Cognitive Reorganization in Response to the Hypnotic Reversal of Attitudinal Affect," Journal of Personality, 28 (1960), 39-63.
- Ryan, Michael J. "A Test of Associations in a Causal Chain Derived from Fishbein's Behavioral Intention Model," unpublished working paper, University of Alabama, 1975.

\_\_\_\_\_, and E.H. Bonfield. "The Fishbein Extended Model and Consumer Behavior," Journal of Consumer Research, 2 (September 1975), 118-37.

Schmidt, Frank L. "Implications of a Measurement Problem for Expectancy Theory Research," Organizational Behavior and Human Performance, 10 (October 1973), 243-51.

\_\_\_\_\_, and Terry C. Wilson. "Expectancy Value Models of Attitude Measurement: A Measurement Problem," Journal of Marketing Research, 12 (August 1975), 366-8.

Thurstone, L.L., and Lyle V. Jones. "The Rational Origin for Measuring Subjective Values," Journal of the American Statistical Association, 52 (1957), 458-71.

Wilson, David T., H. Lee Mathews, and James W. Harvey, "An Empirical Test of the Fishbein Behavioral Intention Model," Journal of Consumer Research, 1 (March 1975), 39-48.

FISHBEIN'S SUBJECTIVE NORM: THEORETICAL  
CONSIDERATIONS AND EMPIRICAL EVIDENCE

Myron Glassman, University of South Carolina  
Nancy Fitzhenry, South Carolina Commission of Alcohol and Drug Abuse

Abstract

A number of questions about Fishbein's behavioral intention model were raised and in some cases empirical findings pertaining to the questions were discussed. The psychological significance of the regression coefficients, the conceptualization of the subjective norm and the specificity of the motivation to comply were dealt with.

History of the Model

In recent years, Fishbein's behavioral intention model has stimulated a great deal of research. (For a review see Fishbein and Ajzen, 1975). While most of the research has been promising, there remain certain methodological and empirical questions about the normative component of the model. As originally formulated (Fishbein, 1967), the model contained two normative components, the personal normative component and the social normative component. Mathematically the model was expressed as:

$$B = BI = (A_B)_{w_0} + (NB_p \times MC_p)_{w_1} + (NB_s MC_s)_{w_2} \quad (1)$$

Where B = overt behavior; BI = behavioral intention;  $A_B$  = the attitude toward performing the behavior;  $NB_p$  = the respondent's personal feelings as to whether or not he should perform the behavior;  $MC_p$  = his personal motivation to comply;  $NB_s$  = the respondent's perception of the social norms pertaining to the performance of the behavior;  $MC_s$  = his motivation to comply with the social norm; and  $w_0, w_1$  and  $w_2$  = empirically determined weights. Since a high correlation was found between the respondent's personal normative belief ( $NB_p$ ) and behavioral intention (BI), the personal normative component was not included in subsequent research, and the social normative component was reformulated to include the respondent's perceptions of the expectations of significant others regarding his performance of the behavior ( $NB_i$ ) and his general motivation to comply with each significant other ( $MC_g$ ). Mathematically the model was represented as:

$$B = BI = (A_B)_{w_1} + (\sum NB_i MC_g)_{w_2} \quad (2)$$

Most recently, the normative component has been viewed in terms of a subjective norm (SN), the respondent's perception of what "others who are important to him" feel that he should do regarding the performance of the behavior. It should be noted that this formulation of the model means that  $\sum NB_i MC_g$  is no longer viewed as a primary determinant of intention. It has been replaced in the predictor equation by SN. Instead, it is now viewed as the determinant of SN. The implications of the change in the role of  $\sum NB_i MC_g$  will be discussed shortly.

The Relative Importance of  $A_B$  and SN

One finding of the studies dealing with Fishbein's behavioral intention model is that there has been a tendency for behavioral intention to be primarily determined by attitudinal factors (Fishbein, 1975, p. 311).

There are two explanations for this finding. The first is that since the relative weights of the two components are expected to vary as a function of the topic and population under investigation, the slight tendency for attitudinal considerations to play a dominant role in determining intention is due to the fact that for the particular topics and populations studies, attitudinal considerations in fact were dominant. The second explanation is that the impact of normative beliefs in determining intention is felt in the attitudinal component. That is, one of the determinants of an individual's favorable attitude toward buying a particular product might be that he believes that Referent X will be pleased.

Another question pertaining to the relative importance of the normative and attitudinal component concerns the psychological meaningfulness of the weights. While the weights have been shown to vary across topics and populations, one might question whether these weights are meaningful or are merely "statistical artifacts." The results of studies pertaining to this question are inconclusive. Two of the studies (Ajzen and Fishbein, 1970 and Ajzen, 1971) used a Prisoner's Dilemma game. In both studies, some subjects were told to treat the other player as a partner (cooperative condition) while others were told to try to do better than the other person (competitive condition). It might be expected that in the cooperative condition normative factors would be dominant while in the competitive condition attitudinal factors would be dominant. As seen in Table 1, the results were encouraging.

TABLE 1  
ATTITUDINAL AND NORMATIVE WEIGHTS IN  
PRISONER'S DILEMMA GAMES

Study	Condition	Regression Coefficients	
		Attitudes	Norms
Ajzen and Fishbein (1970)	Cooperative	.229	.707 <sup>a</sup>
	Competitive	.691 <sup>a</sup>	.327 <sup>a</sup>
Ajzen (1971)	Cooperative	.112	.768 <sup>a</sup>
	Competitive	.541 <sup>a</sup>	.225

<sup>a</sup><sub>p</sub> < .05

Also of relevance is a study conducted by Glassman (1971). In this field study conducted among working class housewives, 127 respondents were asked questions pertaining to their beliefs, attitudes and intentions to buy eight convenience goods which had been selected because of their particular location on the Cohen and Barban (1970) product typology. Four of the products (two brands each of coffee and laundry detergent) were classified as ego-conspicuous products, i.e., they gave the consumer a sense of personal satisfaction (ego involved) and they were in some way visible to others (conspicuousness), while four of the products (two brands each of gasoline and potato chips) were classified physiological-inconspicuous, i.e., the products did nothing more for the consumer than the jobs for

which they were designed (physiological) and the products and/or the results of their use were not visible to others (inconspicuousness). In addition, respondents completed the C.A.D. personality inventory (Cohen, 1967). This instrument classifies individuals according to their interpersonal orientation. Specifically, an individual can be viewed as moving toward people (compliant), against people (aggressive) or away from people (detached). Since perceived social conspicuousness has been called the most general attribute bearing on a product's susceptibility to reference group influence (Bourne, 1963), it was hypothesized that the intention to purchase those products that were classified as being ego-conspicuous would be primarily determined by normative factors while the intention to purchase those products that were classified as being physiological-inconspicuous would be primarily determined by attitudinal factors.

TABLE 2  
REGRESSION COEFFICIENTS FOR EIGHT CONVENIENCE GOODS

	Regression Coefficient		Multiple R
	Attitude	Norms	
a Folgers Coffee	.467 <sup>c</sup>	.186	.608 <sup>c</sup>
a Hills Bros. Coffee	.569 <sup>c</sup>	.207 <sup>c</sup>	.707 <sup>c</sup>
a Tide Detergent	.665 <sup>c</sup>	.105	.742 <sup>c</sup>
a Cheer Detergent	.461 <sup>c</sup>	.133	.529 <sup>c</sup>
b Shell Gasoline	.442 <sup>c</sup>	.271 <sup>c</sup>	.662 <sup>c</sup>
b Standard Gasoline	.500 <sup>c</sup>	.255 <sup>c</sup>	.680 <sup>c</sup>
b Jays Potato Chips	.651 <sup>c</sup>	.120	.724 <sup>c</sup>
b Lays Potato Chips	.540 <sup>c</sup>	.179 <sup>c</sup>	.665 <sup>c</sup>
a Ego-conspicuous product			
b Physiological-inconspicuous product			
c $p < .05$			

It can be seen in Table 2 that irrespective of product classification, attitudinal considerations were the primary determinant of intentions. In fact, for three of the four ego-conspicuous products, the regression coefficient of the normative component was insignificant. However, it was encouraging to note that normative considerations played a significant role in determining intention to purchase gasoline (physiological-inconspicuous). Specifically, subsequent analyses showed that the women felt that their husbands had very definite ideas as to which gasoline should be purchased and there was a strong motivation to comply with this referent. It was also felt that the extent to which a respondent was compliant, aggressive or detached, would influence the amount of importance that she placed on each of the components of the model. Specifically, it was hypothesized that for those women who classified themselves as being compliant, intention would be primarily a function of normative considerations while for those women who were aggressive or detached, attitudinal considerations would be primary. Since 70.6% of the women classified themselves as being primary compliant it was felt that a compliant versus aggressive and detached analysis would be inappropriate. Rather, the sample was divided into three groups on the basis of their compliance score. The results did not lend support for the hypothesis.

#### The Subjective Norm

It will be recalled that the behavioral intention model specifies SN, the generalized normative belief, as the normative predictor of intention. That is,  $\sum NB_i MC_g$  is no longer placed directly in the regression equation as a predictor of BI, but rather is viewed as the determinant of SN. The results of a study conducted by

Fitzhenry and Glassman (1974) support the notion that  $\sum NB_i MC_g$  is the determinant of SN. One hundred and thirty females who were enrolled in an introductory psychology course at the University of Illinois during Fall 1974 answered questions pertaining to the various components of the behavioral intention model for (1) the following topics: taking birth control pills, donating blood, engaging in premarital sex, using amphetamines, buying a car, drawing up a will and using marijuana and (2) the following referents that were chosen based on responses to an elicitation question: parents and friends. Of relevance to the discussion is the form of the following questions:

- SN: Others who are important to me think I should +3: : : : -3:  
I should not perform Behavior X
- MC<sub>g</sub>: In general I want to do 7: : : : :1:  
do not want to do what Referent A thinks I should do
- NB: Referent A would think that my performing Behavior X would be  
Good +3: : : : : -3: Bad

While the measure of NB is not in the traditional form, it is consistent with Fishbein's (1975, p. 315) conceptualization of the variable. As seen in Table 3, the results are consistent with Fishbein's theorizing. In a similar study by Glassman and Birchmore (1974), sixty females who were introductory psychology students at the University of Illinois during the Fall of 1972 were asked to fill out a questionnaire which dealt with (1) various components of the intention model for several contraceptive behaviors: use of the birth control pill, use of a diaphragm, use of the rhythm method, use of an I.U.D., asking their boyfriend/husband to obtain a vasectomy, obtaining contraceptive foam and asking their boyfriend/husband to obtain a condom and (2) the following elicited referents: physician, mother, father, friend, planned parenthood, priest, husband/boyfriend and women's magazines. In addition, motivation to comply was measured at three levels: general (MC<sub>g</sub>), moderate (MC<sub>m</sub>) and specific (MC<sub>s</sub>).

TABLE 3  
THE RELATIONSHIP BETWEEN SN AND  $\sum NB_i MC_g$

Topic	Coefficients
Take "the pill"	.530 <sup>a</sup>
Donate Blood	.153 <sup>a</sup>
Engage in Premarital Sex	.426 <sup>a</sup>
Use Amphetamines	.573 <sup>a</sup>
Draw up a Will	.487 <sup>a</sup>
Use Marijuana	.427 <sup>a</sup>

a  $p < .05$

Of relevance to this discussion is the wording of the following questions:

- SN: Others who are important to me think I should +3: : : : -3: should not perform Behavior X
- NB: Referent X thinks that I should +3: : : : : -3: I should not perform Behavior X
- MC<sub>g</sub>: In general I want to do 7: : : : :1:  
do not want to do what Referent X thinks I should do
- MC<sub>m</sub>: With respect to sexual behavior I want to do 7: : : : :1:  
do not want to do what Referent X thinks I should do
- MC<sub>s</sub>: With respect to matters of birth control I want to do 7: : : : :1:  
do not want to do what Referent X thinks I should do

TABLE 4  
CORRELATIONS BETWEEN SN AND  $\sum NB_i MC$

Formulation	Use Pill	Use Diaphragm	Use Rhythm	Topic	Obtain Vasectomy	Obtain Foam	Obtain Condom
				Use I.U.D.			
SN = $\sum NB_i MC$ general	.707 <sup>a</sup>	-.108	.795 <sup>a</sup>	.907 <sup>a</sup>	.707 <sup>a</sup>	.731 <sup>a</sup>	.493 <sup>a</sup>
SN = $\sum NB_i MC$ moderate	.721 <sup>a</sup>	-.095	.671 <sup>a</sup>	.859 <sup>a</sup>	.664 <sup>a</sup>	.767 <sup>a</sup>	.513 <sup>a</sup>
SN = $\sum NB_i MC$ specific	.750 <sup>a</sup>	.014	.786 <sup>a</sup>	.827 <sup>a</sup>	.750 <sup>a</sup>	.707 <sup>a</sup>	.537 <sup>a</sup>

a p < .05

As seen in Table 4, the results generally support the notion that SN is determined by the sum of the normative beliefs weighted by the general motivation to comply.

While Fishbein normally stresses the importance of specificity, he opts for a general motivation to comply. His argument (Fishbein, 1975, p. 306) is that if MC were measured at a specific level it would merely reflect the normative component's regression coefficient. However, Franch and Raven (1959), the same source that Fishbein uses as a foundation for the motivation to comply suggest that an individual's influence is likely to be topic specific. In fact, given that  $\sum NB_i MC$  is no longer placed in the regression equation that predicts intention, the argument relating a specific motivation to comply ( $MC_G$ ) to the normative regression coefficient seems irrelevant.

In its new role,  $\sum NB_i MC$  can be conceptualized as follows. The respondent, in arriving at his perception of the subjective norm, can be viewed as either consciously or unconsciously taking into account the desires of relevant referents and weighting each referent's desires by the relative importance of the referent's opinions. In this light, the determinants of SN can be viewed in terms of the respondent's perceptions of the desires of "important others" and a set of weights. Within this framework, measuring MC at a general level may be inappropriate in that at best  $MC_G$  can be viewed as a measure of general importance, i.e., if a person generally tends to comply with an individual it can be assumed that the individual is generally important to that person. However, it is very likely that with respect to a specific behavior, this person's opinions may not be important. Similarly, while a person's opinions about a specific behavior may be important, in general they may not be.

While a specific motivation to comply may be more appropriate given the above framework, there is the possibility that questions using an  $MC_G$  format will be responded to in terms of the respondent's intention and his perception of the desires of others. If the respondent intends to perform the behavior and views the referent as thinking that he should perform the behavior, he will be motivated to comply with the referent regarding the performance of the behavior. Likewise, if the respondent intends to perform the behavior and he feels that the referent does not want him to perform it, he will not be motivated to comply with the referent regarding the performance of the behavior.

It would seem that based on the previous discussion both  $MC_G$  and  $MC_S$  are inappropriate measures of the motivation to comply. Returning to Table 4, it appears that from an empirical point of view it makes little difference as to whether the motivation to comply is measured at a general, moderate or specific level.

Another issue pertaining to the motivation to comply is whether or not it should be used as a weight for SN. Again, Fishbein has argued that the motivation to comply with "important others," irrespective of the level of specificity, would merely reflect the regression coefficient of the normative component. Some evidence relating to this question comes from the Glassman and Birchmore (1974) study. Table 5 reports the average multiple regression coefficients for the seven intentions studied.

TABLE 5  
AVERAGE MULTIPLE REGRESSION COEFFICIENTS

Model	Coefficients
BI = $A_B + SN$	.514
BI = $A_B + SNMC_G$	.506
BI = $A_B + SNMC_M$	.514
BI = $A_B + SNMC_S$	.508

It can be seen that the addition of MC irrespective of the level of specificity does nothing to improve the predictability of the model. As an additional note, all multiple regression coefficients were significant except those dealing with the use of the diaphragm. These non-significant findings were due, in part, to the lack of variability in the intention to perform the behavior.

#### Additional Considerations

Given the change in the role of  $\sum NB_i MC_G$  there is a question as to whether the traditional question used to elicit referents, "Where would you go for more information or advice concerning Behavior A?" is appropriate. Within the context of the previously mentioned framework, it would seem that the major interest would be in referents whose opinions concerning the performance of the behavior are important. It would seem best if respondents were now asked, "With respect to your performance of Behavior A, whose opinions are important to you?" While there could be a considerable correspondence between the results of the two elicitation questions it is likely that differences will occur. For example, while a girl may not consider going to her mother for information or advice about using the birth control pill because of possible fear or embarrassment, it is likely that she would know how her mother felt about her use of the pill, even if this information were based solely on inferences, and it is likely that she would consider her mother's opinions as being important. While the first elicitation question would not be expected to evoke the response "mother," the

second one would.

A final point has to do with the way in which the motivation to comply is phrased. Usually respondents are asked whether they "want to do" or "do not want to do" what a particular referent thinks that person should do. If scored from +3 (want to comply) to -3 (do not want to comply) a spite model is assumed. That is, if the respondent feels that Referent A thinks he should not perform the behavior (NB = -3), and he is not motivated to comply with Referent A (MC = -3), Referent A will contribute positively toward his intention to perform the behavior (-3 x -3 = 9). If scored from +6 (want to comply) to 0 (do not want to comply), the fact that the respondent does not want to comply with referent A (MC = 0) who thinks that he should perform the behavior (NB = +3) implies that the wishes of Referent A make no contribution to the respondents intention to perform the behavior (+3 x 0 = 0). Research to date indicates these different methods of scoring have little influence on the regression weights and multiple correlation coefficient.

#### Summary

In general, the findings support the most recent conceptualization Fishbein's behavioral intention model. There is evidence that the regression weights are psychologically meaningful; SN was shown to be a function of  $\sum NB_i MC_i$  irrespective of the specificity of MC; and intention was shown to be predictable from  $A_B$  and the subjective norm, irrespective of whether or not SN was weighted by the motivation to comply.

#### References

- Icek Ajzen, "Attitudinal vs. Normative Messages: An Investigation of the Differential Effects of Persuasive Communications on Behavior," Sociometry, 34(1971), 263-280.
- Icek Ajzen and Martin Fishbein, "The Prediction of Behavior from Attitudinal and Normative Variables," Journal of Experimental Social Psychology, 6(1970), 466-487.
- Frank S. Bourne, "Different Kinds of Decisions and Reference-group Influence,": in P. Bliss (ed.), Marketing and the Behavioral Sciences (Boston: Allyn and Bacon, 1963).
- Joel B. Cohen, "An Interpersonal Orientation to the Study of Consumer Behavior," Journal of Marketing Research, 4(1967), 270-278.
- Joel B. Cohen and Arnold M. Barban, "An Interactive Consumer-Product Typological System: A Progress Report and Partial Evaluation," Working Series in Marketing Research, College of Business Administration, Penn State University, 1970, Paper No. 12.
- Martin Fishbein, "Attitudes and the Prediction of Behavior," in Martin Fishbein, ed. Readings in Attitude Theory and Measurement. New York: John Wiley and Sons, Inc., 1967, 477-492.
- Martin Fishbein and Icek Ajzen, Belief, Attitude, Intention and Behavior, An Introduction to Theory and Research, (Reading, Mass.: Addison-Wesley, 1975).
- Nancy Fitzhenry and Myron Glassman, "The Prediction of Intention," Unpublished Manuscript, University of Illinois, 1974.
- John French and Bertram Raven, "The Bases of Social Power," in D. Cartwright, ed., Studies in Social Power (Ann Arbor: University of Michigan Press, 1967).
- Myron Glassman, "The Effects of Personality and Demographic Factors on the Formation of Attitudes Towards Convenience Goods," Unpublished masters thesis, University of Illinois, 1971.
- Myron Glassman and Melinda Birchmore, "The Relationship Between Subjective Norms and Normative Beliefs," Unpublished Manuscript, University of Illinois, 1974.

TOWARD A VECTOR MODEL OF INTENTIONS

Olli T. Ahtola, University of Florida

Abstract

This paper takes a critical look at Fishbein's model of behavioral intentions. It argues that the model as conceptualized and operationalized by Fishbein and his associates contains some inconsistencies, unacceptable assumptions, and unexplained components. A new model is developed which proposes a solution to these weaknesses.

Fishbein's Model of Intentions

The so called "extended Fishbein model" was developed to predict and explain a specific behavioral act in a well defined situation. In a given situation, a person is assumed to form a specific behavioral intention which directly determines his subsequent overt behavior. According to the theory, there are two major factors that determine behavioral intentions: a personal or "attitudinal" factor and a social or "normative" factor. These two factors in the theory are given empirical weights. Symbolically, the central equation in the theory can be presented as follows (Fishbein and Ajzen, 1975, p. 301):

$$B - I = (A_B) w_1 + (SN)w_2 \quad (1)$$

where:

- B ≡ the behavior (in a specified situation)
- I ≡ the intention to perform behavior B
- A<sub>B</sub> ≡ the attitude toward performing behavior B
- SN ≡ the subjective norm, i.e., the influence of the social environment on behavior B
- w<sub>1</sub> and w<sub>2</sub> are empirically determined weights

According to the theory, the person's attitude toward performing the behavior (A<sub>B</sub>) is a function of the perceived consequences of performing that behavior and of the person's evaluation of those consequences. The functional form is proposed to be (Fishbein and Ajzen, 1975, p. 301):

$$A_B = \sum_{i=1}^n b_i e_i \quad (2)$$

where:

- A<sub>B</sub> ≡ the attitude toward performing behavior B
- b<sub>i</sub> ≡ the strength of belief that performing behavior B leads to consequence or outcome i
- e<sub>i</sub> ≡ the evaluation of outcome i
- n ≡ number of salient beliefs

Furthermore, according to the theory, the subjective norm (SN) is determined by the sum of the perceived expectations of specific referent individuals and/or groups weighted by the individual's "motivation to comply" with those expectations.

Algebraically (Fishbein and Ajzen, 1975, p. 302):

$$SN = \sum_{i=1}^n b_i m_i \quad (3)$$

where:

- SN ≡ the subjective norm

- b<sub>i</sub> ≡ the normative belief (i.e., the person's belief that referent group or individual i thinks he should or should not perform behavior B)
- m<sub>i</sub> ≡ the motivation to comply with referent i
- n ≡ number of relevant referents

Empirical Weights

The behavioral intention is proposed to be a weighted sum of the two factors - "attitudinal" and "normative". Standardized multiple regression coefficients have been used to estimate the weights. Empirical evidence shows that the weights vary with the behavior, with the situation, and with individual differences. These weights are proposed to represent the individual's relative importance of the two factors (Fishbein and Ajzen, 1975). Why these importance weights have never been tried to measure directly and independently is not explained. Also, how is it possible to least squares estimate these weights if they vary with the behavior and with individual differences? If we want to least squares estimate the weights, at the minimum, there should be a theory which stipulates what kinds of individuals do not vary in the weights. Then the weights could be properly estimated for these individuals. Only constants (parameters), not variables, can be estimated through least squares fitting.

Furthermore, it is a well-known fact that if the predictor variables are correlated the precision of estimation falls so that it becomes very difficult, if not impossible, to distangle the relative influences of the predictor variables. Also, estimates of coefficients become very sensitive to particular sets of sample data, and the addition of a few more observations can sometimes produce drastic shifts in the coefficient. Empirical evidence shows that A<sub>B</sub> and SN (or  $\sum b_i m_i$ ) tend to be correlated and, consequently, the above problems are likely. Given the nature of the model and the assumptions in multiple regression, it seems that the regression analysis is quite improper method to estimate the weights, especially when our purpose is to explain the intentions and not just to predict them, i.e., we are interested in the values of the weights and our sole purpose is not just to maximize R<sup>2</sup>.

Personal Attitude

The first factor, i.e., the "attitudinal" factor, is a conceptually clear measure of how much the individual likes or dislikes to perform the behavior. Often, however, this factor has been measured in terms of the individual's attitude toward the behavior (e.g., Ajzen and Fishbein, 1969) and not in terms of the individual's attitude toward he himself performing the behavior. The latter, of course, is the conceptually correct procedure. Fishbein makes a distinction between attitude toward an object and attitude toward a particular behavior (with respect to the object). It would seem logical that he made a distinction also between attitude toward the behavior in general, and attitude toward a particular actor (like the subject himself) performing the behavior.

Even though A<sub>B</sub> variable in itself is conceptually clear and easy to operationalize, the model of its determi-



nants (2) has been argued to have several problems (Ahtola, 1973, 1975). However, because these problems have been discussed in detail elsewhere (Ahtola, 1973, 1975) they will not be discussed in this paper.

#### Normative Influence

"The second or normative component of the theory, SN, deals with the influence of the social environment on behavior. The subjective norm is the person's perception that most people who are important to him think he should or should not perform the behavior in question." (Fishbein and Ajzen, 1975, p. 302)

The above statement indicates that Fishbein believes that there exists in an individual's cognitive structure a unidimensional concept, subjective norm, which is structurally similar to attitude. It should be noted that this unidimensional variable was not proposed in his earlier writings (e.g., Fishbein, 1972). In his earlier writings Fishbein used normative beliefs and motivation to comply as first order determinants of intention without any intermediary variable, such as the subjective norm.

The question arises whether people really perceive any collective "important others" in normal circumstances. Important others are usually such things as spouse, father, mother, employer, friends, etc. It seems obvious that in many cases, perhaps in most cases, the opinions of these referent individuals or groups conflict or at least vary in magnitude. If this is the case, do referents, such as spouse, employer, and friends, get cognitively combined into a concept "important others"? It would seem plausible that several different persons in a group, such as a social club, where members hold relatively homogenous norms about expected behavior of their members, will be cognitively combined into a collective group, such as "the members of my Rotary Club". But it seems very unlikely that the wife's opinion would be cognitively combined with the club members' opinion to form a generalized opinion of "important others". It is argued here that very seldom does there exist such a thing as generalized subjective norm in people's cognitive structure, except perhaps when the individual is contemplating some abnormal or asocial behavior which is uniformly sanctioned by the whole society.

Furthermore, if Fishbein's model of the determinants of SN (3) is correct it seems inconsistent to conceptualize (and operationalize) the SN as "the person's perception that most people who are important to him think he should or should not perform the behavior in question" (Fishbein and Ajzen, 1975, p. 302). This arises from the fact that the motivation to comply can be negative as well as positive, i.e., the individual may be motivated not to do what the referent wants. Let's take a hypothetical example in which each relevant referent of a person thinks he should perform the behavior but he is strongly motivated not to comply with their wishes. Under those circumstances  $E_{b_i}$  gives a highly negative score. If, however, the SN is measured directly as suggested (Fishbein and Ajzen, 1975, p. 314), i.e., in terms of whether "most people who are important to me think I should --- I should not perform behavior X", one definitely would expect a highly positive response. These are totally opposite results. This inconsistency results from the fact that the "subjective norm" measure does not seem to incorporate the "motivation to comply" influence. The SN measure is probably some kind of average  $b_i$  score over the relevant referents, which, as indicated earlier, most likely has very little, if any, cognitive meaning to the individual.

Turning to the determinants of the social norm, one finds additional problems. These problems, however, are such that they do not conflict with the above argument about the questionable conceptual meaning of the social norm, i.e., while evaluating the determinants of SN it will not be assumed that SN is correctly conceptualized. In other words, there are internal problems in this submodel.

The normative belief,  $b_i$ , is conceptualized as the person's belief that a relevant reference group or individual thinks he should or should not perform the behavior. There are clearly two conceptually independent variables in this conceptualization. One is the strength of the subject's belief and the other is the direction and polarity of the referent's opinion. Fishbein, however, measures  $b_i$  unidimensionally. Perhaps the operationalization of  $b_i$  indicates which variable is the one conceptualized relevant to the model.

At least three different operationalizations can be found in the literature. These are:

1. (Ajzen and Fishbein, 1969)

Referent X thinks I should perform the behavior  
probable \_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_ improbable

2. (Fishbein and Ajzen, 1975)

Referent X thinks  
I should \_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_ should not  
perform the behavior

3. (Ajzen and Fishbein, 1969)

Referent X expects me to perform the behavior  
probable \_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_ improbable

The above three measures are not conceptually parallel. The first scale measures the strength of the subject's belief. If the subject marks anything but the extreme left category, he is indicating that there is a possibility that the referent is indifferent or against the behavior. The second scale, on the other hand, (especially if the standard Semantic Differentials Scale instructions are given to the subject) seems to ask the subject's perception of the content (direction and polarity) of the referent's opinion, i.e., how much the referent is for or against the behavior. One must draw conclusion that the model is not conceptualized to discriminate between the perceived polarity (i.e., the content) and the uncertainty (i.e., the strength) of the belief. These two aspects of the normative belief are logically different and it seems very plausible to argue that they are also "psycho-logically" different. This is because they need not even co-vary. For example, it is quite conceivable that a person may think that referent X is most likely very much against the behavior but the person is somewhat uncertain about it. At the same time, the person may think that referent Y is slightly unfavorable and the person is certain about it (e.g., because the referent Y has told him what he feels, while the person may be just guessing the referent X's opinion).

It seems very likely that both the uncertainty about the belief and the content of the belief (especially the polarity) affect behavioral intentions, and consequently both of them should be incorporated in the model.

The problem with the third scale is that it may not measure at all what the referent is perceived to think the subject should do. It may measure what the subject thinks the referent thinks he will do. This is the

basic denotative meaning of the scale. The subject may know that his wife expects him to get drunk next Saturday night because that is what he always does. She still probably does not think that he should get drunk.

The second component in the normative influence sub-model, i.e., the "motivation to comply", is suggested to be measured in terms of how much the subject, in general, wants or wants not to do what the referent thinks he should do. There seems to be considerable uncertainty about the exact meaning of this component. Should it be independent of the referent's particular demands, or should it be specific to the particular behavior or behavioral domain under consideration, or should it be defined as the subject's general motivation to comply with the referent? The last conceptualization is advocated by Fishbein (Fishbein and Ajzen, 1975). The theoretical grounds for this choice are not referred to, but the guess of this author is that the grounds are to make the "motivation to comply" component independent of the "normative belief" component. This would be very convenient for predictive purposes; but is it conceptually sound? The subject might want to comply with his wife's strong wishes but could care less about her less polarized wishes. On the other hand, the subject may form a reactance to his father's strong demands, so that he does not comply if he feels high pressure, while if the pressure is lower he is happy to comply. What is argued here is that the "motivation to comply" is not independent of the content of the normative belief.

French and Raven (1959) and Kelman (1961) have convincingly argued that social influence of a referent increases with that referent's power to reward or punish the person (compliance), with the person's liking of, or desire to, identify with the referent (identification), with the referent's perceived expertise (internalization), and with the extent to which it is legitimate for the referent to make demands of the person. If the purpose of the motivation of comply component is to tap all the above influences, its operationalization as a unidimensional single measure seems utterly optimistic. As mentioned earlier, Fishbein suggests that the component be operationalized in terms of how much the subject wants or wants not to go along with the referent's opinion. This scale probably measures the identification influence quite well but it may not get at the compliance influence ("I don't want to do what my father thinks I should do, but I am still going to do as he wants because otherwise he is going to spank me"), or the expertise influence ("I really dislike my doctor and his advice, but I still do what he tells me to do because he is a real expert") or legitimate right influence ("I really dislike to send money to my no-good son at the college, but I'll do it anyway because I got money from my father and that is a tradition and responsibility in our family").

#### Toward a Vector Model of Intentions

The purpose here is to suggest what properties and what kinds of functional relationships the model of behavioral intentions should possess if the above criticism of the Fishbein model is accepted.

To understand and predict a person's intention to perform a specified act, we must know how much he would like to perform the act, i.e., what his personal attitude toward his performing of the act is. However, before deciding whether to perform the act or not he may want to consider what others would like him to do. That is, we must know what he perceives each of the relevant referents' attitude toward him performing the

act to be. In other words, the subject's perception as to how much each referent wants him or wants him not to perform the act. In many cases, there may be uncertainty in the subject's mind about the polarity or perhaps even the direction of some of the referents' opinions.

#### Properties of Perception and Belief

People seldom perceive things along continuous dimensions but instead locate them in terms of categories on these dimensions. For example, some attitudes may be perceived "very favorable" while some others only "fairly favorable" or "slightly favorable". That is, the attitude dimension has been partitioned into categories, and the attitudes are described in terms of these categories, not in terms of some point on a continuous dimension. The number of categories perceived or utilized seems to differ somewhat among individuals (Osgood, Suci, and Tannenbaum, 1957). Seven categories have been found to be very common. Let's assume that the subject is using seven attitude categories, which he has labeled "very favorable", "fairly favorable", "slightly favorable", "indifferent", "slightly unfavorable", "fairly unfavorable", and "very unfavorable".

Each category is a concept which together form the population of attitude concepts which the subject may associate with the referent (i.e., perceive the referent to have). Each referent elicits one or more of these attitude concepts (i.e., attitude categories). It is quite conceivable that under certain circumstances, the same referent elicits two or more of these attitude categories. It seems reasonable to assume that the most likely attitude category to the referent is elicited first (Fishbein, 1967). Then the other possible categories to the referent are elicited in order of their likelihoods. This takes place when the subject is not sure what the referent's attitude is.

For example, the subject has the belief that his mother most likely is "very unfavorable" toward him performing the act, but he is not quite sure and he thinks that his mother may be "fairly unfavorable" or perhaps only "slightly unfavorable". He knows, that his mother is not neutral or favorable.

In summary, each belief statement has the content and the strength. The content in the above example is the direction and the polarity of the referent's perceived attitude and the strength is the likelihood or probability the subject assigns to the referent having the attitude with that content.

#### Motivation To Be Influenced

To understand and predict the effect of the perceived opinion of a referent we must know what the subject's motivation to be influenced by the referent's attitude is.

The subject may, and most likely does, have a different motivation to be influenced to differently polarized attitudes of the referent. The motivation may or may not be monotonically related to the polarity of the referent's perceived attitude. The subject may be motivated not to be influenced if the referent's attitude is very polarized (reactance), while he may be influenced positively by the less polarized wishes. The subject, at the same time, may have zero motivation to another referent's mild attitudes, while, when the attitudes are perceived to be polarized, he is motivated to go along. Father and wife of the subject might be such referents. The subject may be irritated by his father's extreme wishes and shows his independence by not complying, while he complies with his father's less extreme wishes because he accepts them as legitimate. On the other

hand, the subject may not care at all about his wife's small wishes and takes his wife's wishes into consideration only when he perceives them to be really strong.

In operationalizing the motivation to be influenced, at least two scales seem necessary. These are the measures of 1. how much the subject wants or wants not to do as the referent wants him to do given the polarization of the referent's attitude, and 2. how much the subject feels he should or should not do as the referent wants him to do given the polarization of the referent's attitude. These two influences should be summative.

#### Influence of a Referent

The motivations to be influenced are viewed as summative. Furthermore, the amount of behavioral response tendency available for summation is an absolute function of the strength of belief. It seems reasonable to argue that if the subject is sure what the referent's wish is, all the corresponding motivation is available, i.e., the corresponding motivation to be influenced has the weight 1. On the other hand, if the subject perceives no chance that a given attitude category represents the referent's opinion, none of the corresponding motivation to be influenced is available, i.e., the weight is 0. Various levels of uncertainties have weights between 0 and 1. The more likely the category for the referent the higher the weight. Because every referent must have some attitude (note, indifference is an attitude) the sum of the subjective probabilities assigned to different attitude categories for each referent should add up to 1.

Algebraically, this may be expressed as follows:

$$M_r = \sum_{j=1}^g b_{rj} m_{rj} \quad (4)$$

where:

- $M_r$   $\equiv$  motivation to be influenced by referent r (with respect to behavior B)
- $b_{rj}$   $\equiv$  the strength of belief that referent r has attitude j
- $m_{rj}$   $\equiv$  motivation to be influenced by referent r's attitude j
- g = number of attitude categories

or in vector notations:

$$M_r = \underline{b}'_r \underline{m}_r \quad (5)$$

where:

- $\underline{b}'_r$   $\equiv$  row vector of probabilities of referent r's association with the attitude categories (as perceived by the subject)
- $\underline{m}_r$   $\equiv$  column vector of the subject's motivation to be influenced by referent r's attitude categories

#### Intention

The influences of the various relevant referents are viewed as summative. The influence of the personal attitude and the influences of the various relevant referents are viewed as summative. Because the personal attitude and the influences of the referents are measured along somewhat different dimensions and using different scales, a scale conversion constant (which does not have any cognitive meaning) is needed.

In summary, the proposed model can be represented symbolically as follows:

$$I_B = A_B + \beta \sum_{r=1}^n \underline{b}'_r \underline{m}_r \quad (6)$$

where:

- $I_B$   $\equiv$  an individual's intention to perform behavior B
- $A_B$   $\equiv$  his personal attitude toward performing B
- $\beta$   $\equiv$  scale conversion constant
- n  $\equiv$  number of relevant referents
- $\underline{b}'_r$  and  $\underline{m}_r$  are the same as in (5)

#### References

- Olli T. Ahtola, "An Investigation of Cognitive Structure within Expectancy-Value Response Models," Unpublished doctoral dissertation (University of Illinois, 1973).
- Olli T. Ahtola, "The Vector Model of Preferences: An Alternative to the Fishbein Model," Journal of Marketing Research, 12 (February, 1975), 52-59.
- Icek Ajzen and Martin Fishbein, "The Prediction of Behavioral Intentions in a Choice Situation," Journal of Experimental Social Psychology, 5 (1969), 400-416.
- Martin Fishbein, "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.
- Martin Fishbein, "The Search for Attitudinal-Behavioral Consistency," In J. Cohen (ed.), Behavioral Science Foundations of Consumer Behavior (New York: The Free Press, 1972), pp. 245-252.
- Martin Fishbein and Icek Ajzen, Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research (Reading: Addison-Wesley, 1975).
- John P.R. French, Jr. and Bertram H. Raven, "The Bases of Social Power," In D. Cartwright (ed.), Studies in Social Power (Ann Arbor: University of Michigan Press, 1959), pp. 150-167.
- Herbert C. Kelman, "Processes of Opinion Change," Public Opinion Quarterly, 25 (1961), 57-78.
- Charles E. Osgood, George J. Suci, and Percy H. Tannenbaum, The Measurement of Meaning (Urbana: University of Illinois Press, 1957).

THE NATURE OF SALIENT OUTCOMES AND REFERENTS  
IN THE EXTENDED MODEL

Michael J. Ryan, The University of Alabama  
Michael J. Etzel, Utah State University

Abstract

The specification of outcomes and referents, although a crucial aspect in the development of the Extended Model, has received little attention. This paper discusses saliency and empirically investigates the elicitation technique. The issues raised by the findings indicate that a good deal of research is needed to investigate both the nature of saliency itself and methods for ascertaining salient outcomes and referents.

Introduction

The specification of attributes has been considered a weak part of multi-attribute composition models (Hughes, 1974).<sup>1</sup> Fishbein and Rosenberg, who are generally credited with providing the structural origins of these models, have been cited for providing little guidance in this manner (Wilkie and Pessemier, 1973). Consequently, consumer researchers have developed (Pessemier and Wilkie, 1974) and criticized (Kerman, 1974) methods for attribute specification. Although these developments may prove useful, it seems appropriate to examine the methods promulgated by the originators of these models.

A perusal of Fishbein's earlier writings (e.g., Fishbein, 1963) reveals that he determined salient concepts for the A<sub>0</sub> model with an elicitation procedure based on free association techniques developed by Maltzman (Maltzman, 1955; Maltzman, Bogartz, and Breger, 1958). This method continues to be used in the extended model by Fishbein and his associates (e.g., Jaccard and Davidson, 1972) and has been suggested (Fishbein, 1971) and used successfully in advertising research applications (Cowling, 1973a).

This paper discusses the elicitation technique and empirically examines specific issues related to its use. Behavioral outcomes and referents determined by way of elicitation from similar groups were compared across two brands in the same product class; the order in which outcomes and referents were most frequently elicited was compared to individual importance rankings; and the individual importance rankings were analyzed in terms of consistency by group and brand.

Elicitation and Saliency

It is indigenous to the extended model that it be based on salient outcomes and referents. An outcome is considered salient if it serves as a determinant of the

<sup>1</sup>Although the operational procedures and model structures addressed in this paper are basically the same, there is a major difference between expectancy value models that incorporate product attributes and those, such as the extended model, that utilize behavioral outcomes. The reader unfamiliar with this distinction is referred to the review provided by Ryan and Bonfield (1975). A useful distinction and empirical comparison among common social psychological and consumer research conceptualizations of attitude models is also furnished by Mazis, Ahtola, and Klippel (1975).

attitude toward a specific behavioral act and if the individual associates the outcome with the act. Although Fishbein has not been explicit about the source of salient referents, the structural similarity between the attitudinal and normative components suggests that a referent would be salient if it serves as a determinant of the subjective norm and if the individual associated the referent with the act. The position taken in this paper is that saliency issues, usually referred to hereafter in terms of outcomes, also hold for referents. Also, the relationship of outcomes with the act is specified according to variable combination rules exhibited by the model. However, structure and other model variables will be mentioned as little as possible in order to avoid confusion and allow a clearer focus on the issue of saliency.

Fishbein (1971) has stated that the only way to determine salient outcomes is a simple elicitation procedure in which individuals are asked what outcomes they associate with a specific behavioral act. It is possible to elicit outcomes associated with but not determiners of the attitude. However, based on information processing research (Miller, 1956), Fishbein believes that only 5 to 9 outcomes are salient (Fishbein, 1967, 1971) and that these salient outcomes are the first in order of elicitation. In fact, the notion of elicitation and saliency are intertwined.

Saliency refers to the fact that the respondent is aware of or conscious of the attribute, that it's on the "tip of his tongue." In other words, it has a high probability of being elicited by the respondent. (Fishbein, 1971: p. 313).

The execution of the elicitation technique consists of asking respondents from a defined group a series of non-directive questions, such as:

- Q. What comes into your mind when you think about buying Crest toothpaste?
- Q. When I say to you buying Crest toothpaste what do you think of?
- Q. Is there anyone you know who might like or dislike you to buy Crest toothpaste?

The simplicity of this technique will not appeal to those who believe consumers cannot divulge such associations in so straightforward a manner (Dichter, 1960; Martineau, 1957). Yet, in a different context, a direct elicitation technique worked as well when compared to more complicated methods and produced results superior to those obtained with indirect questioning (Alpert, 1971).

Although a detailed criticism of alternative procedures has been furnished elsewhere (Cowling, 1973; Fishbein, 1971), some of these criticisms will be mentioned briefly. First, factor analysis is not deemed appropriate since salient items, while being conceptually distinct, are not always expected to score differently on an index of strength, magnitude, etc. Factor analysis merely identifies sets of items that are highly intercorrelated. For example, white teeth and decay prevention are measures of some relative characteristic, and if correlated, would indicate one factor although they are obviously conceptually distinct outcomes. Al-

though this example is extreme and many would argue that a skilled interpretation of a factor loadings matrix would not produce this result, there may be more subtle distinctions made in the minds of consumers that would not be distinguishable with factor analysis. Also, if the notion that consumers are limited information processors is accepted, there seems to be little to gain in reducing an already small number of dimensions. The theory assumes that outcomes are combined, according to their respective beliefs and evaluations, in an unweighted additive manner.<sup>2</sup> Only salient items are included in the model. Therefore, all included outcomes and referents are determiners of attitude and social influences respectively. The relative determining power of these few salient items is indicated only in the combination of their respective beliefs and evaluations.

#### Some Issues and Problems

Use of the elicitation procedure has produced good results in correlational studies (Cowling, 1973); Bonfield and Ryan, 1975, Ryan, 1975) and has been shown to be more effective than the use of predetermined lists (Mazis, *et al.*, 1975). Yet, there are some issues and problems that need to be explored. This research examined four such issues.

First, there is the issue of whether or not the elicitation procedure provided generalizable outcomes salient to different groups and brands. Given the wide academic acceptance of the marketing concept and the resulting segmentation strategies, it seems consumer researchers would avoid applying the same set of attributes or outcomes to different groups. For example, using outcomes elicited from college students in a study of housewife purchasing behavior conflicts with all we know about behavioral differences in marketing. Different brand positioning strategies would also suggest that outcomes may be brand rather than product specific. Consequently, this research elicited outcomes for two differently positioned brands across two independent samples similar in life style but geographically separated. It was expected, based on Cowling (1973, 1973a), that elicited outcomes would be similar for the two groups but different for each brand.

Second, the question of whether the procedure will provide outcomes relevant to group members or merely provide idiosyncratic items needs to be directly addressed.<sup>3</sup> Although the majority of Fishbein's current work is group oriented (Fishbein and Ajzen, 1975), he has argued that the best estimate of attitude is obtained from a cognitive structure based on individual subject's own elicited concepts (Kaplan and Fishbein, 1969). Also, the majority of success in applying elicitation procedures have used correlational analysis which can be misleading in terms of model fitting (Birnbaum, 1973). Consequently, outcomes were tabulated by frequency and order of mention in order to ascertain if outcomes could be identified that were common to the majority of group members.

Third, the relative importance of the outcomes, aside from the information derived from belief and evalua-

<sup>2</sup>Research questioning these assumptions has been reviewed by Ryan and Bonfield (1975) and more recent evidence questioning the additivity assumption has been put forth by Troutman and Shanteau (1975).

<sup>3</sup>Wilkie and Pessimier (1973) have described this problem succinctly and suggested individual level analyses across brands as a possible solution.

tion statements, is of interest. Yet, the unweighted composition rules do not allow a direct assessment of importance. Fishbein (1971, 1975) has argued that concepts that are more important have more polarized belief and evaluation statements. He believes that this explains why importance weights, when added as a third cognitive structure variable, do not increase prediction. That is, their effects are already present in the belief and evaluation statements. Although all strongly held beliefs need not be based on salient outcomes, all salient outcomes involve strong beliefs and elicitation is the only way to ascertain saliency (Fishbein, 1971). From this, it seems reasonable to conclude there may be a direct relationship between importance and order of elicitation. This is not to suggest an importance variable as a replacement or addition to the model; rather, it may be useful in terms of parsimony or verification. If order of elicitation indicated outcome importance this information by itself could be useful before construction of belief and evaluation statements or as a cross check on belief and evaluation measurement procedures. Consequently, this research examined the relationship between the order in which outcomes were elicited and individual rankings of outcome importance.

Fourth, given that the above relationship is supported, the use of individual importance rankings suffers from the same limitations in applying individually elicited concepts to all group members. Namely, are the rankings idiosyncratic or is there consistency among group members? Consequently, the individual rank orderings were examined for consistency.

#### Method

Subjects consisted of two convenience samples of undergraduate students enrolled in upper division business courses. The first group, containing 97 subjects, was located at The University of Alabama and the second group, containing 121 subjects, was located at the University of Kentucky. These groups were chosen to represent similar lifestyles although they were expected to be slightly different due to separate regional influences. The product class chosen was toothpaste and the brands were Crest and Ultra Brite. The product class was chosen since it has been a popular subject of multiattribute model researchers who have sometimes used predetermined attribute lists across brands and widely varying sample groups. The brands were chosen to represent different positioning strategies; Crest being positioned as a decay preventative and Ultra Brite as a cosmetic. All data were collected during the Spring of 1975.

Subjects were asked to write their answers to a series of questions similar to those shown above pertaining to each brand. Responses were analyzed and outcomes and referents were tabulated according to total frequency of mentions and frequency of mentions by rank order. Outcomes and referents were then ordered according to frequency of mentions. Finally, each subject rank ordered the outcomes in terms of importance.

Kendall's Tau statistic was used to compare the rank order of elicitation according to frequency of mentions with each subject's independent ranking. Kendall's Coefficient of Concordance was employed to examine the consistency of the independent rank orderings.

#### Findings and Discussion

The frequencies of outcome and referent mentions for the Alabama sample are shown in Table 1 and for the Kentucky sample in Table 2. The total frequency

Table 1

Outcome and Referent Total Frequency and Frequency  
Of Mentions By Rank Order For Alabama Sample  
(n = 97)

Outcomes	Total Freq.	Crest Frequency by Rank Order					
		1	2	3	4	5	6
Decay Prevention	73	34	22	6	3	6	2
Flavor	41	8	10	14	7	1	1
Color	17	1	5	3	7	1	0
Leading Brand	10	2	1	1	4	2	0
<b>Referents</b>							
Dentist	51	28	15	6	1	0	0
Family	50	24	20	5	0	1	0
Close Friends	17	1	9	4	3	0	0
Outcomes	Total Freq.	Ultra Brite Frequency by Rank Order					
		1	2	3	4	5	6
Sex Appeal	43	27	13	0	1	2	0
White Teeth	32	15	10	5	1	0	1
Favor	26	4	10	9	2	1	0
Fresh Breath	12	0	4	5	2	1	0
<b>Referents</b>							
Dentist	22	10	9	1	2	0	0
Family	19	11	5	6	5	0	0
Close Friends	18	10	4	4	0	0	0

Table 2

Outcome and Referent Total Frequency and Frequency  
Of Mentions By Rank Order For Kentucky Sample  
(n = 121)

Outcomes	Total Freq.	Crest Frequency by Rank Order				
		1	2	3	4	5
Decay Prevention	119	58	34	20	5	2
Flavor	45	10	10	14	8	3
Dental Assn. Endorsement	36	8	11	9	2	6
Color	13	0	7	6	0	0
<b>Referents</b>						
Family	75	36	34	5	0	0
Dentist	53	35	12	5	1	0
Close Friends	16	3	7	6	0	0
Outcomes	Total Freq.	Ultra Brite Frequency by Rank Order				
		1	2	3	4	5
Sex Appeal	90	70	14	5	1	0
White Teeth	49	19	22	6	2	0
Flavor	26	4	14	6	2	0
Abrasiveness	21	6	8	4	1	2
<b>Referents</b>						
Close Friends	43	32	10	1	0	0
Family	20	12	3	4	1	0
Dentist	15	7	6	2	0	0

of mentions and the frequency by rank order were consistent for both samples and brands. That is, those outcomes and referents mentioned first were also the most often mentioned. It is interesting to note that consistent with Cowling's (1973 a) findings, few items were mentioned by more than 50% of the respondents. Although no rules have been established concerning the number of mentions necessary before an elicited outcome is usable, it does seem inappropriate to ask respondents to rate outcomes when the majority of the

outcomes are not salient for the majority of the sample. The good results that have been obtained in studies using such items may be the result of halo effects (Beckwith and Lehmann, 1975). Cowling (1973) has suggested using only the individual's own elicited beliefs although Kaplan and Fishbein (1969) could only conclude that estimates of attitude based on a subject's own elicited beliefs are at least as accurate as estimates based on standard sets of beliefs. At any rate, an idiosyncratic attitude model has limited use to consumer researchers primarily occupied with the understanding of group behavior.

The problem of group saliency becomes more important in realistic settings since the groups under examination are likely to be more heterogenous than those examined in this study. It may be possible, as suggested by Cowling (1973), to start with such a group using common outcomes as a basis for segmentation. On the other hand, a larger sample in the present study may have allowed the identification of subgroups with a large percentage of commonly mentioned outcomes.

A rule of thumb stating that only items mentioned by the majority of respondents be included in the model seems reasonable. Yet, for the groups under study this rule would allow only one outcome and two referents for the Alabama Crest model, no outcomes or referents for the Alabama Ultra Brite model, and one outcome and referent for the Kentucky Crest and Ultra Brite model. Is it realistic to suppose that so few or possibly no outcomes and referents are determiners of attitude and social influence respectively? Possibly so. The notion that 5 to 9 outcomes are expected to be salient is based on studies involving beliefs and attitudes likely to be central to the individual such as racial attitudes (Fishbein, 1967) and attitudes toward birth control (Jaccard and Davidson, 1972). It seems unlikely that toothpaste brand purchase was a central activity for the subjects employed in this study and thus there may have been a very simple, if any, cognitive structure present. In agreement with Nakanishi and Bettman (1974), the cognitive complexity assumed in using numerous outcomes and referents may be an overkill for a low involvement product class such as toothpaste.<sup>4</sup> In the interest of both parsimony and greater understanding more work is needed to explore models based on single or few outcomes. It also seems important for those testing or exploring the extended model to provide readers of their research with a detailed rationale for the inclusion of outcomes and referents in the model. This would allow research consumers to judge the saliency of outcomes regardless of correlational results.

The outcomes and referents shown in Table 1 and 2 are compared across brands and samples in Table 3. In both the Kentucky and Alabama sample, only one outcome was common to both brands. This finding supports the rather commonly held notion that consumers may associate quite different sets of outcomes with different brands. The implication for brand preference attitude research is important since it indicates that respondents may

<sup>4</sup>Nakanishi and Bettman (1974) found that the addition of outcomes to cognitive structure is an order determined by an independent importance ranking did not significantly improve  $A_{act}$  prediction. However, a pre-determined outcome set developed for a different sample was applied across seven toothpaste brands. Consequently, it can be argued that this result was due to a failure to apply brand and group specific outcomes.

Table 3

## Outcomes and Referents

		Alabama Sample	
		Crest	Ultra Brite
Outcomes	Decay prevention		Sex appeal
	Flavor		White teeth
	Color		Flavor
	Lending brand		Fresh breath
Referents		Dentist	Dentist
		Family	Family
		Close friends	Close friends

		Kentucky Sample	
		Crest	Ultra Brite
Outcomes	Decay prevention		Sex appeal
	Flavor		White teeth
	Dental Association endorsement		Flavor
	Color		Abrasiveness
Referents		Family	Close Friends
		Dentist	Family
		Close friends	Dentist

Table 4

Summary of Individual Correlations  
Between Order of Elicitation and Independent Rank Orderings

TAU* Values	Outcomes			
	Crest**		Ultra Brite**	
	Alabama Sample	Kentucky Sample	Alabama Sample	Kentucky Sample
1.00	7	39	0	2
.67	45	65	9	10
.33	30	9	6	21
.00	5	5	24	29
-.33	2	2	35	30
-.67	1	1	16	25
1.00	7	0	7	4
n	97	121	97	121

TAU* Values	Referents			
	Crest**		Ultra Brite**	
	Alabama Sample	Kentucky Sample	Alabama Sample	Kentucky Sample
1.00	50**	20	29	29
.39	35	74	19	18
-.33	5	21	16	6
1.00	7	6	33	8
n	97	121	97	121

\*Based on rank test for each respondent  
 \*\*Frequency of occurrence for individual respondent tau values

base their cognitive structures on different outcomes for different brands. This phenomenon may not manifest itself in response to belief or evaluative rating scales based on non-salient outcomes due to a number of extraneous influences such as demand artifacts (Sawyer, 1975). Again, this points to the need for researchers to explicitly state their rationale for including outcomes in the model. It also indicates that the application of the same outcomes across brands may not provide an adequate test of the extended model. In fact, the lack of common outcomes would obviate strategies based on identifying outcomes with which consumers differentiate brands. It may be that a gestalt approach, considering the entire salient set, is more appropriate if a segment does not have common outcomes across brands.

There were few differences in outcomes across the two samples for the same brand and no differences in referents across brands or samples except for the order of elicitation. Since there was reason to expect the samples to be homogenous, this gives some indication that the elicitation procedure is reliable.

The findings from the comparison of order of elicitation and individual rank orderings of outcomes and referents are summarized in Table 4. There was a direct relationship between the majority of independent rankings and elicited order of Crest outcomes for both samples. The Crest findings, indicating that the first few elicited outcomes were also the most important for the groups, tends to support the previous suspicion that only one or two outcomes may be salient. However, the Ultra Brite findings showed no preponderance of direct relationships. It appeared that respondents

may rank outcomes mentioned by few group members just as important as those mentioned many times. This indicates that importance and saliency are different. These different findings across the two brands are not readily explainable. Yet, the similar results with two independent samples suggests these findings may be reliable. The order of elicitation of Ultra Brite outcomes was the same for the first three outcomes for both samples (Tables 1 and 2). It may be that respondents were reluctant to rank sex appeal as more important than other more rational appearing outcomes when in fact this would have been an accurate response. Again, this may indicate demand artifacts or halo effects. Perhaps the methods suggested by Sawyer (1975) would be useful in investigating this phenomenon. At any rate, whatever the intended use of outcome importance rankings, these findings indicate they may be misleading if applied across brands.

The referent findings were similar. There was a direct relationship between order of referent elicitation and independent importance rankings for the majority of respondents in the Crest model across both groups and for the Ultra Brite Kentucky sample. There were small differences in the frequency of referent mentions for the Alabama Crest sample (Table 1) indicating that these items were equally salient for group members whereas the elicited mentions produced more clearly defined ordering in the other three instances. Also, although the same referents were elicited, the elicited order was different across both brands in the Kentucky sample and the Crest Alabama sample. This finding suggests that the relative importance of referents is indicated by the order of elicitation, and may be brand or group specific.

The findings investigating the consistency of the importance rankings are presented in Table 5. The Crest

outcome rankings were moderately consistent across both samples whereas the Ultra Brite outcome rankings showed

Table 5

Concordance\* Among Ranking of Outcomes and Referents

	Crest		Ultra Brite	
	Alabama Sample	Kentucky Sample	Alabama Sample	Kentucky Sample
Outcomes	.46	.36	.03	.04
Referents	.65	.67	.36	.34

\*Kendall's coefficient of concordance

no consistency across both samples. This finding supports the previous indication that Ultra Brite outcome importance rankings are not meaningful for the two groups. On the other hand, the moderate consistency in Crest importance rankings indicates that in spite of the direct relationship between order of outcome elicitation and importance rankings, there was a good deal of heterogeneity among the individual rankings. Taken together, these findings suggest that outcome importance rankings, since they appear to be idiosyncratic, have little use when applied to groups.

The Crest referent rankings were highly consistent whereas the Ultra Brite rankings were moderately consistent for both groups. When considered with the previous findings, only the Crest referent importance ranks appeared to have meaning in terms of group analysis since they confirmed importance as indicated by order of elicitation and were consistent among respondents in both samples.

Summary and Conclusion

The findings from the independent rankings were largely inconsistent. In general, the consistency of outcome rankings varied across brands and groups as well as did the relationship between independent rankings and order of elicitation as ascertained by frequency of mentions. In addition, there was some indication that respondent importance rankings may be subject to artifact influences. Consequently, independent rankings appeared to be of little value. On the other hand, the suggestion that few or no common outcomes may be held salient by members of homogenous groups is of considerable importance since it implies that commonly used formulations of the model may be overly complex for low involvement products.

Based on this study, the following suggestions are made:

1. The appropriate number of elicited outcomes or referents germane to a target group may be a function of the centrality of the beliefs being studied. A reasonable rule of thumb may be to include only items mentioned by more than 50% of the group members.
2. Outcomes and referents used in studies should always be generated from the group to be studied and should be situation specific. Using items generated from other audiences or for other products or brands may prove misleading.
3. The order of item elicitation may be more useful than importance rankings. Importance ranks may be subject to social desirability bias or other yet to be explained factors that undermine their usefulness.

4. A number of studies using varied groups, products, and brands, and comparative methods are needed to further address the questions raised here. In the meantime, it is advisable to use the elicitation technique in the care and recognition of its potential weaknesses.

References

- Mark I. Alpert, "Identification of Determinant Attributes: A Comparison of Methods," Journal of Marketing Research, 8 (May, 1971), 184-191.
- Neil E. Beckwith and Donald R. Lehmann, "The Importance of Halo Effects in Multi-Attitude Models," Journal of Marketing Research, 12 (August, 1975), 265-275.
- Michael H. Birnbaum, "The Devil Rides Again: Correlation as an Index of Fit," Psychological Bulletin, 79 (April, 1973), 239-242.
- E. H. Bonfield and Michael J. Ryan, "A Practical Application in the Fishbein Extended Model in Ordering Promotional Strategy," unpublished working paper. College of Commerce and Business, The University of Alabama, 1975.
- Anthony B. Cowling, "Use of Elicitation Technique for Producing Dimensions of Brand Choice," Proceedings, British Market Research Society, 1973a, 139-156.
- Anthony C. Cowling, "Determining and Influencing Consumer Purchase Decisions," European Research, 1 (January, 1973), 26-31, 43.
- Ernest Dichter, The Strategy of Desire. Garden City, N. J.: Doubleday and Co., 1960.
- Martin Fishbein, "A Behavior Theory Approach to the Relations Between Beliefs About an Object and the Attitude Toward the Object," in Martin Fishbein, ed., Readings in Attitude Theory and Measurement. New York: John Wiley and Sons, Inc., 1967, 389-400.
- Martin Fishbein, "A Consideration of Beliefs, and Their Role in Attitude Measurement," in Martin Fishbein, ed., Readings in Attitude Theory and Measurement. New York: John Wiley and Sons, 1967, 257-266.
- Martin Fishbein, "An Investigation of the Relationships Between Beliefs About an Object and the Attitude Toward that Object," Human Relations, 15 (August, 1963), 322-340.
- Martin Fishbein, "Some Comments on the Use of 'Models' in Advertising Research," Proceedings. European Society of Market Research, Seminar on Translating Advanced Advertising Theories into Research Reality, 1971, 297-318.
- Martin Fishbein, "Attitude, Attitude Change, and Behavior: A Theoretical Overview," in Philip Levine, Ed., Attitude Research Bridges the Atlantic. Chicago: American Marketing Association, 1975, 3-16.
- Martin Fishbein and Icek Ajzen, Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Reading, Mass.: Addison-Wesley, 1975.
- G. David Hughes, "Trends in the Development of Multi-Attribute Models of Central Information Processing," in G. David Hughes and Michael L. Ray, eds., Buyer/Consumer Information Processing. Chapel Hill: The University of North Carolina Press, 1974, 277-287.



James J. Jaccard and Andrew R. Davidson, "Toward an Understanding of Family Planning Behaviors: An Initial Investigation," Journal of Applied Social Psychology, 2 (July-September, 1972), 228-235.

Kalman J. Kaplan and Martin Fishbein, "The Sources of Beliefs, Their Saliency, and Prediction of Attitude," The Journal of Social Psychology, 78 (June, 1969), 63-74.

Jerome B. Kernan, "Comments on the Role of Summated Attribute Models in the Central Processing of Information," in G. David Hughes and Michael L. Ray, eds., Buyer/Consumer Information Processing. Chapel Hill: The University of North Carolina Press, 1974, 353-355.

Irving Maltzman, "Thinking: From A Behavioristic Point of View," Psychological Review, 62 (July, 1955), 275-286.

Irving Maltzman, William Bogartz, and Louis Breger, "A Procedure for Increasing Word Association Originality and Its Transfer Effects," Journal of Experimental Psychology, 56 (1958), 392-398.

Pierre Martineau, Motivation in Advertising. New York: McGraw-Hill, 1957.

Michael B. Mazis, Olli T. Ahtola, and R. Eugene Klippel, "A Comparison of Four Multi-Attribute Models in the Prediction of Consumer Attitudes," Journal of Consumer Research, 2 (June, 1975), 38-52.

G. A. Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," Psychological Review, 62, (1956), 81-97.

Masao Nakanishi and James R. Bettmen, "Attitude Models Revisited: An Individual Level Analysis," Journal of Consumer Research, 1 (December, 1974), 16-21.

Edgar A. Pessemier and William L. Wilkie, "Multi-Attribute Choice Theory - A Review and Analysis," in G. David Hughes and Michael L. Ray, eds., Buyer/Consumer Information Processing. Chapel Hill: The University of North Carolina Press, 1974, 288-330.

Michael J. Ryan, "A Test of Associations in a Causal Chain Derived from Fishbein's Behavioral Intention Model," unpublished working paper. College of Commerce and Business, The University of Alabama, 1975.

Michael J. Ryan and E. H. Bonfield, "The Fishbein Extended Model and Consumer Behavior," Journal of Consumer Research, 2 (September, 1975), 118-136.

Alan G. Sawyer, "Demand Artifacts in Laboratory Experiments in Consumer Research," Journal of Consumer Research, 1 (March, 1975), 20-30.

C. Michael Troutman and James Shanteau, "Do Consumers Judge Product Quality by Adding or Averaging Attribute Information," Kansas State University Human Information Processing Institute Report #75-4, July, 1975. Also presented at The American Psychological Association Meetings, Chicago, 1975.

William L. Wilkie and Edgar A. Pessemier, "Issues in Marketing's Use of Multi-Attribute Attitude Models," Journal of Marketing Research, 10 (November, 1973), 428-441.

## EXTENDING THE EXTENDED MODEL: SOME COMMENTS

Martin Fishbein, University of Illinois, Urbana-Champaign

### Abstract

This paper reviews the previous papers presented in this session. An attempt is made to clarify some misconceptions about "the extended model."

It's always rewarding to find that other people are interested in your research, and it's even more rewarding when some of those other people turn out to be your ex-students. Since there is perhaps even more normative pressure to engage in discourse with ex-students than other colleagues, I'd like to try to answer some of the questions they've raised and, at the same time, try to clarify a few points about the so-called "extended model". Since a great many different issues have been raised in their papers, and since several of the papers focused on similar questions, I'd like to proceed by dealing with the issues, rather than by commenting on each paper individually.

Although nobody objected to the notion of an attitude toward an action, there were several questions raised about the appropriateness of the concept of "subjective norm". Further, there were some questions about how each of these more global concepts were related to their hypothetical determinants. Finally, there were also questions about the relative weights of these two "global" concepts as determinants of intentions and behaviors.

Generally speaking, in its most recent version, the model states that a person's intention to perform any behavior is a function of (1) his attitude toward performing that behavior ( $A_B$ ) and (2) his subjective norm concerning that behavior (SN). The model also suggests that these two variables may take on different weights in determining different intentions. That is, some intentions may be entirely under attitudinal control, other intentions may be entirely under normative control, and still other intentions may be influenced by both attitudinal and normative considerations. The relative weights of these two components as determinants of an intention are expected to vary as a function of the type of intentions being considered and individual difference variables. Further, the model suggests that, like any other attitude, the attitude toward the behavior can be viewed as a function of the person's salient beliefs (in this case about performing the behavior) and the evaluative aspects of those beliefs ( $A_B \approx \sum b_e$ ). Similarly, the subjective norm is viewed as a function of normative beliefs and the person's motivation to comply with relevant referents ( $SN \approx \sum b_m$ ). These ideas have been outlined in the earlier papers presented in this session. The most complete statement of the model appears in Chapter 7 of the book I recently wrote with Icek Ajzen (Fishbein & Ajzen, 1975).

The Attitudinal Component. Given the vast amount of literature that has been generated by the expectancy-value model (and various adaptations or modifications of it), I was somewhat surprised (and I must admit delighted), to find that relatively little attention was directed at this component in today's papers. Although there seems to be general agreement that attitude can be assessed by locating the attitude

object on a bipolar evaluative dimension, Olli did raise one question about the conceptualization of the  $A_B$  variable. He suggested that the appropriate attitude object should be the person's own performance of the behavior rather than "performing the behavior" per se. I agree with Olli, but I do not think that this distinction is always necessary. That is, empirically, it turns out that in most cases, a person's attitude toward "performing" a given behavior and his attitude toward "my performing" a given behavior are very highly correlated. In these cases, there is probably no need for the distinction. However, if you have any reason to suspect that these two attitudes are different, then the appropriate attitude would be the subject's attitude toward his or her performing the behavior. Icek and I (Ajzen & Fishbein, 1975) made this point in a recent review of the attitude-behavior literature. We came across two studies that used an expectancy-value model to predict women's contraceptive behaviors (Insko et al, 1970; Kothandapani, 1971), and one of these studies was much more successful than the other. As Olli suggested, this difference appeared to be primarily due to the attitude being measured. That is, in the Insko et al study the predictor variable was a woman's attitude toward "using a contraceptive" or while in the Kothandapani study her attitude toward "my using a contraceptive", was assessed. The latter measure yielded significantly better prediction.

Turning to the determinants of attitude, two very different types of questions have been raised. Ryan and Etzel were concerned with the question of how one identified salient beliefs, and Lutz raised some specific questions about the measurement of beliefs, as well as the combination rule underlying the expectancy-value model. Further, as I'm sure you all know, in an earlier paper (Ahtola, 1975) Olli also raised questions about belief measurement; in particular, the problem of distinguishing between belief strength and belief content.

There is unfortunately, relatively little I can say about the paper by Ryan and Etzel. Although I agree that the problem of identifying salient beliefs is a crucial one, I don't really see that their paper contributes very much to our current body of knowledge. What these investigators have shown is that when beliefs are elicited about different brands in the same product class, different beliefs may be salient for the different brands, and further, that the order (or frequency) of elicitation is not related to some independent estimate of the importance of the elicited attribute (or attribute dimension). Neither of these findings are new or exciting. Anthony Cowling (1973), as well as other investigators in Great Britain, has previously shown that different beliefs, or if you like, different attributes or attribute dimensions, are associated with different brands in the same product field. Similarly, Kalman Kaplan and I (Kaplan & Fishbein, 1969) provided evidence that judgments of attribute importance were not related to order of elicitation or frequency of elicitation of an attribute. In our recent book, however, Icek Ajzen and I (Fishbein & Ajzen, 1975) suggested that an attribute's importance might be reflected in the strength of a person's belief that the product

has the attribute, and/or in his evaluation of that attribute. That is, it seems reasonable to assume that in many cases the more important the attribute, the more information a person may have about it, and thus the more certain he should be that the product has (or does not have) the attribute in question. Similarly, in many cases, the more important the attribute, the more likely the evaluation of the attribute will be extreme (or polarized). However, people can hold strong beliefs about attributes they view as unimportant, and they may evaluate unimportant attributes very favorably or unfavorably. Thus, we would not expect a strong relation between either belief strength and importance or evaluation and importance. However, the importance of a given attribute may be reflected in the absolute value of the  $b_{e_i}$  score associated with the attribute in question. Since either  $b_i$  or  $e_i$  (or both) may be at an extreme when the attribute is important, the absolute value of the  $b_{e_i}$  score should be higher when the attribute is important than when the attribute is unimportant. Please note that if this hypothesis is correct, it would provide strong evidence that "importance" is taken into account by the expectancy-value formulation, and we could do away with importance once and for all. On the other hand, if the hypothesis is not correct, it does not mean importance should be incorporated in the expectancy-value formulation. We already know that this will not improve prediction (see e.g., Fishbein & Ajzen, 1975, Ch. 6). Further, underlying my use of the expectancy-value formulation is the assumption that a person's salient beliefs are the "important" beliefs since they determine the attitude. It's for this reason that it is necessary to first identify salient beliefs if one wants to understand the determinants of a given attitude. I would like to see Ryan and Etzel reanalyze their data in the light of these considerations.

Although the Ryan and Etzel paper has not contributed anything new to our body of knowledge, I don't want to simply dismiss it offhand. I do think it's very important to recognize that the beliefs that determine a person's attitude toward one brand in a product class may be very different from the beliefs that determine his attitude toward another brand in the same product class. Although I have not thought this out in any great detail, findings like those of Ryan and Etzel may have important implications for models and approaches that assume that choice behavior is determined by a comparison of two or more brands along a common set of "relevant" or "important" dimensions. I might buy A rather than B because I think buying A is better than buying B but if the reasons underlying these two attitudes are very different, how meaningful is it to ask respondents to compare the two brands (or the two behaviors) on a set of common scales? I don't know the answer to this question but I think it is one possible direction that Ryan and Etzel could take.

While the questions raised by Ryan and Etzel are not central to the model, Rich and Olli have raised questions about the expectancy-value formulation per se. Rich asks whether the measure of belief should be a measure of probability or a measure of association; Olli made a similar distinction in his earlier work by pointing to the difference between belief content and belief strength. While I do feel that this distinction is a valid and important one, I also feel that there is a great deal of confusion surrounding its relevance to an expectancy-value formulation. Olli's point is that the

content of beliefs can differ; one person may believe that "X is slightly sweet" while another may believe that "X is quite sweet". Similarly, one person may believe that "Performing X leads to Y" while another may believe that "Performing X blocks Y". As Olli pointed out, these are different beliefs and a person might assign a high probability to the statement "X is slightly sweet" but a low probability to the statement "X is quite sweet". Thus, Olli has argued that an appropriate model must take both belief strength and belief content into account. Rich bases his argument on the fact that Rosenberg and I have used different scales to measure belief. That is, Rich points out that Rosenberg (1956) asked subjects whether a policy would "lead to" or "block" the attainment of a given outcome, while I have usually asked respondents whether it is "probable" or "improbable" that a behavior will lead to a given outcome. I don't know how many times I've said this before, but let me say it again -- both Rosenberg and I agree that at a conceptual as well as an operation level, the central equations in our models (i.e.,  $\sum b_{e_i}$  versus  $\sum I_i V_i$ ) are identical. This is not to say that our theories are identical, for there are some very major differences between us, but the central equations and their operationalizations are identical. That is, both Rosenberg and I would agree that if we knew that a person believed that "X is quite sweet", we would both want to know the strength (probability) of this belief, and the evaluation of "quite sweet". Similarly, if we knew that a person believed that "Performing behavior X blocks Y", we would both want to know the strength (probability) of this belief and the evaluation of "blocking Y". To argue that our models are different because Rosenberg once used a "leads to - Block" scale and I usually use a "likely - unlikely" scale, ignores our theories and reifies operations into conceptual definitions. While I agree that wording and scale formats are important, and while it's true that I may get somewhat different results if I use a "leads to - blocks" scale than if I use a "probable - improbable" scale, it's also true that I'd get somewhat different results if I used a "likely - unlikely" scale, but this would not mean that I had redefined the concept of belief.

I think the real question is not one of whether the belief measure should be an associative or probability measure, but rather a question of why both Rosenberg and I have used a bipolar belief measure if we're in agreement that belief strength or probability is what we're after. The answer to this question is relatively straightforward, and it's largely pragmatic.

I think it would be fair to say that both Rosenberg and I assumed that if a person believed that a given object did not have a negative attribute or if a given policy did not lead to (or blocked) a negative consequence this would imply a favorable attitude toward the object or policy. If beliefs were scored on a straight probability basis, (i.e., from 0 to 1), or on any unipolar scale, they could not reflect this double-negative. We both chose to measure our beliefs on a bipolar scale so that we could capture this double-negative property.

In retrospect, I have come to realize that the double-negative assumption is not necessary. More specifically, viewing beliefs as only associative rather than as associative or disassociative, forces one to reconsider the definition of the outcome or attribute associated with the object of belief. For example,

consider the belief "X is not sweet". I can either view this as a belief that disassociates X and "sweet" or as a belief that associates X and "not sweet". Similarly, the belief that "X blocks Y" can either be viewed as a disassociation between X and "Y" or an association between X and "blocks Y". Viewing all beliefs as associative eliminates the need for a double-negative assumption, since the contribution of the belief to attitude will always be reflected in the evaluation of the associated attribute. Thus, as I indicated above, if the content of a person's belief is known, there is no question about how that belief should be measured -- I would want to measure the strength of that belief on a unipolar probability scale and the evaluative aspect of that belief on a bipolar evaluative scale. These measurements will tell me whether the belief is contributing positively or negatively to the overall attitude. For example, if someone believes that "Performing X blocks Y" and he thinks that "blocking Y" is "bad", I know this belief contributes negatively to this attitude. Similarly, if I know a person believes X leads to Y and that he positively evaluates Y, I know that this belief is contributing positively to his overall attitude. The problem comes when I don't know the content of the person's belief, i.e., when I don't know whether he believes that "performing X leads to or blocks Y". Consider the person described above who believes "X blocks Y" -- if I simply asked him to indicate how probable it was that "X leads to Y" he would indicate a low probability. Further, if I asked him to evaluate Y, he would evaluate it positively. Thus, if I scored the probability judgment from zero to one, I would infer that this belief was contributing positively to his attitude (e.g.,  $.20 \times +2 = .40$ ). However, we have already seen that his belief is really that X blocks Y and he thinks that blocking Y is bad. Thus, this belief should contribute negatively to his attitude. By treating the belief dimension as a bipolar dimension (e.g., by subtracting .5 from the obtained probability) the negativity is captured (i.e.,  $.2 - .5 = -.3 \times +2 = -.6$ ). What this procedure does not do is take into account the fact that the person's evaluation of (blocking Y) may not be the opposite of his evaluation of (Y). In our recent book, Icek Ajzen and I (Fishbein & Ajzen, 1975) discussed these problems in detail, and we suggested that when one could not assume symmetry in evaluation, the appropriate procedure would be to measure both of these beliefs. Olli has carried this even further by suggesting that we ask the respondent to provide us with all of his beliefs along the attribute (or outcome) dimension. That is, we should ask people how probable it is that "Behavior X leads to Y", that "Behavior X is unrelated to Y", that "Behavior X blocks Y", etc. He wants to identify all possible positions on a given dimension, measure belief strength with respect to each position, and then measure the evaluation of each of these possible outcomes. That is, we should also ask them to evaluate "leads to Y", "does not lead to Y", "blocks Y", etc.

To return to the main point however, the problem is not in knowing how to define belief, but in arriving at a measurement procedure that best represents the theory. Both Rosenberg and I relied on bipolar measures of belief and evaluation. As Rich's work shows, this does correspond quite well to the cognitive algebra used by subjects. To summarize briefly then, if you don't know the content of a person's belief and if you can only measure one belief per attribute dimension (or if you only want to measure one belief) the most appropriate thing to do

is to treat the belief dimension as bipolar. However, if you know the content of a person's belief, then you should use a unipolar measure. Whether the benefits obtained from asking more than one question per attribute dimension (i.e., whether following a procedure such as the one Olli suggested) will outweigh the costs (in subject time and cooperation as well as in instrumentation) is an empirical question that can only be answered by additional research.

Rich raised one other question about the expectancy-value model -- he reopened the old problem of adding versus averaging. While time does not permit a complete discussion of this problem, let me just say that although the Bettman, Capon and Lutz (1975) studies are an improvement upon earlier studies of adding vs. averaging, they still fail to provide an adequate test for resolving the adding-averaging controversy. At this point in time I'm still convinced that the relevant data supports an additive model, and while I am not willing to completely rule out the possibility that there may be averaging on some occasions or in some situations, I will stick to adding until valid data suggests otherwise. For a more complete discussion of this problem, see Fishbein and Ajzen (1975, Ch. 6).

The Normative Component. Turning to the second component, let me first say that I am pleased that most of the questions raised today concern this part of the model since I am more than willing to admit that it is here that most work is needed. However, with the exception of Mike Glassman's paper, I'm afraid that most of the questions raised are more likely to be harmful (i.e., lead to fruitless research) than helpful (i.e., solve some basic problems). In order to respond to these questions, I think it's first necessary to provide some perspective. It should be recalled that until very recently, the normative component was comprised of normative beliefs (i.e., beliefs that specific referents think I should or should not engage in the behavior in question) weighted by motivation to comply with the referents. This  $\sum b_m$  value was thus viewed as an immediate determinant of the intention. More recently, however, the model has been revised to include a general normative concept which has been defined as a "subjective norm". That is, just as beliefs and their evaluative aspects are seen as contributing to a more general attitudinal concept, it was assumed that normative beliefs and motivations to comply contribute to a more general normative concept. Since it seemed reasonable to assume that people would have little or no motivation to comply with unimportant others, I felt that one way to express or define this concept was in terms of "most people who are important to me", and thus we predicted that this subjective norm, i.e., this belief that "most people who are important to me think I should (or should not) engage in the behavior" would influence intentions and would be determined by  $\sum b_m^j$ . Olli questions the validity of a concept such as a "subjective norm". He states that he can't see people thinking in terms of "most people who are important to me". Rich, on the other hand, does not seem to object to a general theoretical concept like "subjective norm", but he questions whether defining it in terms of "important others" really takes the normative prescription of negative referents into account. Olli makes a similar point. Thus, while both Olli and Rich seem willing to accept the notion that people hold normative beliefs (i.e., beliefs that specific referents think they should or should not perform the behavior, and while they both seem

willing to accept the notion that people are differentially motivated to comply with a given referent, Olli doesn't think that this information can be summarized by a single global concept (whatever it's called). Rich thinks it can be represented at a general level, perhaps even by a concept of "subjective norm", but he questions my particular formulation or definition of the "subjective norm".

There is very little I can say in response to Olli, except to point out that respondents have no trouble answering questions about "most important others". Further, I agree with Olli that even if I believed that all referents thought I should perform a behavior ( $b_i = +1$ ), and if I were motivated not to comply with all the referents ( $m_j = -1$ ), I probably would still say that "most people who are important to me think I should engage in the behavior". That is, even though the  $\sum b_i m_j$  score would be negative, the way I measure "subjective norm" would probably produce a positive score. However, I cannot conceive of a situation where I would be motivated not to comply with any referent. That is, although this may be a problem in the abstract, I don't think it presents a practical problem. What we're really talking about here is an empirical question, namely, how well does  $\sum b_i m_j$  predict SN. Mike Glassman's paper provides some initial evidence that  $\sum b_j m_j$  actually does quite well.

More important, I think Mike is using a viable approach to attack some of the real problems with the normative component and its determinants. In most of my early work with the model, motivation to comply contributed little, if anything, to the understanding of intentions. At first, it seemed to me that this might have to do with the level at which motivation to comply was measured. Mike's studies measuring motivation to comply at three different levels of specificity is an important first step in untangling this concept. Unfortunately, his data suggest that the level at which one measures motivation to comply really doesn't seem to matter. While I still think this problem has to be resolved, I now think the major problem resides in the instrument we have used to measure motivation to comply. Theoretically, I have assumed that if a person believed that a given referent didn't want him to engage in some behavior and if the person were motivated not to comply with that referent, this would increase his intention to perform the behavior. Thus, as in the case of belief and evaluation we have measured normative beliefs and motivation to comply on bipolar scales (e.g., from +3 to -3) in order to capture this double negativity. As Mike pointed out, it's only recently that we have come to realize that this bipolar scoring is probably inappropriate for a scale that ranges from "I want to" to "I want not to do what referent X thinks I should do." This wording is clearly unipolar, and consistent with this, when we went back and rescored the motivation to comply measure from 1 to 7 or 0 to 6, we found that motivation did contribute something to prediction, although its contribution was still relatively small. Rich's data also demonstrates that, at least in its present form, the motivation to comply measure should have been scored in a unipolar fashion since people treat it unipolarly in their cognitive algebra. Unfortunately, this does not capture the underlying theoretical assumption, and I personally feel that as a next step we should measure motivation to comply on a true bipolar scale. A scale like "I want to do - I want to do the opposite of what referent X thinks I should do" would be more

in keeping with our underlying assumptions and may give motivation to comply an opportunity to contribute more to prediction. If nothing else, this type of scale will at least lead to a multiplicative cognitive algebra. I will return to this problem below.

First however, let me turn to the more specific suggestions that were raised. Although Olli suggested doing away with a general normative concept, he did want to retain the sum of normative beliefs weighted by motivation to comply. However, he again tries to make a distinction between belief content and belief strength, and he proposes the use of a vector model for this component as well as for the attitudinal component. Although the vector model may have some use on the attitudinal side, I do not think that this will be the case on the normative side. The reason for this is that despite Olli's elaborate argument, I do not think there is a problem of belief content when it comes to normative beliefs. The distinction between belief strength and belief content is important on the attitudinal side, for even if I know a person's belief that "X leads to Y", I cannot predict his belief that "X blocks Y". I can, however, predict his belief that "X does not lead to Y." That is,  $P(X) = 1 - P(\bar{X})$ . On the normative side, all we are interested in is whether the person believes that a given referent thinks he should  $[P(X)]$  or should not  $[P(\bar{X})]$  engage in a particular behavior. Further qualifications such as "Referent A thinks I always should, or Referent A thinks I sometimes should engage in the behavior" will be reflected in the probability measure  $P(X)$ . What I'm trying to point out is that I will obtain essentially the same information whether I use a "should-should not" scale or a "probable-improbable" scale, as long as I treat the probability scale as bipolar. Olli's argument that my two different measures imply two different concepts leads him to make the same mistake Rich made earlier -- he's reifying operations into concepts without paying attention to theory.

Olli was correct however, when he pointed out that, in at least one study, I did use an operation which did, in fact, define a different concept. Asking someone to indicate what another person expects him to do, is not the same as asking him what that other person thinks he should do. Mistakes like this are serious ones, and here differences in measurement procedures do change the meaning of the concept. But notice I used the word mistakes, and I would like to take this opportunity to make a general point. We all make mistakes, and in trying to develop the model I did, once, make the mistake of measuring normative beliefs in terms of expectations. But I realized this was a mistake, I realized that expectations are not the same as shoulds. It has been over seven years since that study was conducted, and in every paper since then, I have defined normative beliefs in terms of "should". Similarly, as I pointed out earlier, many people have tried to make an issue out of the fact that Rosenberg measured instrumentality on a "block-attain" scale while I measure belief on a probability scale. But Rosenberg used that scale in 1956; in his later work, he, too, referred to beliefs in terms of probability. I think that if someone is going to criticize another person's work, or compare the work of two people, the least the critic can do is consider the person's current work. One purpose of research is to enable us to learn and modify our thinking. I don't feel we should be accused of inconsistency because of early mistakes that have been corrected in more recent work.

To return to the problem at hand, however, Olli had one other suggestion or comment -- namely he argued that the motivation to comply measure was not tapping all of the bases of social influence and he suggested some alternative measures. Olli may be right but at this stage he's merely speculating, and I think it's a serious mistake for anyone to suggest changes in someone else's measurement procedures without first providing empirical evidence that such changes are warranted. As I said earlier, I too think that motivation to comply has been badly measured, but I think that one is much more likely to resolve problems by doing the kind of work that Mike Glassman is doing than by introducing unsupported speculations into the literature. Unfortunately, I think this same criticism applies to many of Rich's suggestions. Since Rich has made several different suggestions, let me try to discuss them one by one.

First, I think it's worth noting that both Rich and Olli essentially objected to the definition of the subjective norm because they felt it did not capture situations in which  $m_j$  takes on negative values. That is, Olli argued that the way I have defined SN could produce a positive value in those cases where the referents thought the person should perform the behavior but the person was not (or was negatively) motivated to comply with the referents (i.e., where  $b_j \times m_j$  is negative). Rich pointed to the opposite problem since he feels that the direct measure of SN might produce a negative value in those cases where the referents thought the person should not engage in the behavior but the person was motivated not to comply with the referents (i.e., where  $b_j \times m_j$  is positive). Although I can understand their theoretical arguments, I do think it's highly unlikely that we'll find very many instances where people are not motivated or are negatively motivated to comply with referents they view as important. That is, it seems reasonable to assume that the more important the referent is to the person, the more he will be motivated to comply with that referent. If this is the case, defining the subjective norm in terms of "most people who are important to me" will in fact be quite representative of the  $\sum b_j m_j$  equation.

Further, it's interesting to note that, at this stage, Rich's own data invalidate his objection. That is, as I pointed out above, the objection is only valid if the measure of motivation to comply can take on negative values. Rich's data, as well as Mike Glassman's, point out that in all our research to date, motivation to comply has been measured as a unipolar construct. Rich, however, obviously does believe that motivation to comply can take on negative values. Indeed, he suggested that the subjective norm would be better defined in terms of a global motivation to comply. That is, by assuming that  $b_j$  (normative beliefs) was analogous to  $b_i$  (instrumental beliefs) and  $m_j$  was analogous to  $e_i$ , Rich arrived at a purely motivational description of the normative component and he suggested that SN could be measured by a scale such as: "With respect to performing behavior X, I want to -- I want not to do what other people think I should do." Unfortunately, as Rich himself points out, although he views this concept as bipolar, it is not directional with respect to the behavior. While a scale such as I want to -- I want to do the opposite" will better capture the bipolarity of the motivational concept, it will still not be directional with respect to the behavior, and despite his claim to the contrary, I really don't see how this

nondirectional concept could contribute to prediction of a directional intention.

Further, I think it may be worth noting that had Rich viewed  $m_j$  as analogous to  $b_j$  and  $b_j$  as analogous to  $e_j$ , he would have argued that the subjective norm should be a general concept reflecting what others think we should or should not do. That is he would have arrived at essentially the same conception of subjective norm as the one I proposed. Further, he would have arrived at this conception even though motivation to comply could take on negative values. What I'm trying to point out is that the definition of SN in terms of most important others is in fact compatible with the  $\sum b_j m_j$  formulation even if  $m_j$  is truly bipolar. But conceptual compatibility does not mean it will be correct empirically. Unfortunately, we will not be able to resolve this question until motivation to comply is measured on a truly bipolar scale.

Rich's alternative suggestion for improving the measure of the subjective norm was that it might be better defined as a "social attitude". This suggestion is just plain wrong. Dividing beliefs about performing a given behavior into two arbitrary categories will not solve any problems. In fact, what a procedure of this type will do is provide two different estimates of exactly the same attitude. Thus, rather than reducing the multi-collinearity between the attitudinal and normative components, this procedure will actually increase it. More important, such a procedure would eliminate any consideration of normative pressure. Since Rich would merely have two different estimates of the person's attitude toward performing the behavior he would have no normative component at all. On the basis of a great deal of previous research in many different areas, I can assure you that, despite the fact that different types of beliefs may be used to estimate the attitude, the two estimates will be very highly correlated.

Perhaps at this point I should mention that at one time, I too felt that the second component might be attitudinal and, like Rich, I thought in terms of a social attitude. That is, I thought that since the first component was the person's own attitude toward performing the behavior, the second component could be his perception of the attitudes of relevant others. What this would mean is that in addition to asking a respondent to indicate his own beliefs about performing the behavior we could also ask him to provide us with his perception of others' beliefs about performing the behavior. Further, we could ask for his own evaluation of the association outcome as well as his perception of the way others would evaluate the related outcome. Fortunately, before spending a great deal of time on this idea, Icek Ajzen and I (Ajzen & Fishbein, 1972) looked at the relation between a person's normative belief (i.e., his belief that a given referent thought he should or should not engage in the behavior) and his perception of the referent's attitude toward engaging in the behavior. The results indicated that although the perception of the referent's attitude was related to the normative belief, it was not equivalent to the normative belief. That is, although people may use their perceptions of a referent's attitude as one basis for inferring a normative belief, there appear to be many other bases for the inference. Although I may believe that a referent positively evaluates the performance of a given act, I may still believe that he thinks I should not

engage in the act. Since perceived attitudes of others did not predict normative beliefs we did not feel it would be fruitful to view the normative component as a perceived attitude. Mike Glassman's paper provides additional evidence for the inequality of normative beliefs and perceived attitudes of others. If you look at Tables 2 and 3 in Mike's paper, you will see rather large differences in the size of his obtained correlations between  $\xi_{b,m}$  and SN. Note, however, that in Table 2,  $b_i$  was incorrectly measured in terms of the perceived attitude of the referent, while in Table 3,  $b_i$  was correctly measured, i.e., subjects indicated their belief that a given referent thought they should or should not engage in the behavior.

At this stage in the development of the theory, I do feel that there is some general normative concept that corresponds to the general attitudinal concept. Further, although I am not convinced that the best way to measure this is by asking whether "most important others think I should or should not perform the behavior", I think that this question still comes closest to capturing the meaning of the concept. Further, there is some fairly good evidence that when the normative concept is defined in this way, it is predictable from  $\xi_{b,m}$ . Unfortunately, this may largely be a function of the fact that  $m_i$  has been poorly measured. That is, as I tried to point out above, the notion of important others very adequately reflects  $\xi_{b,m}$  when  $m_i$  is treated in a unipolar fashion. Although I don't know what will happen when  $m_i$  is measured in a manner consistent with the theory, i.e., when  $m_i$  is measured in a true bipolar fashion, I am predicting that SN will continue to be well defined in terms of "important others." Needless to say, I see the motivation to comply concept as the weakest link in the theory. While a bipolar measure may help, this still will not resolve the question of the level at which the concept should be measured. I am convinced that it would be inappropriate to measure motivation to comply in terms of the subject's desire to perform the specific behavior prescribed the referent. Further, although Ick and I have most recently opted for a general measure of motivation to comply with the referent, I am now leaning toward the middle level, namely motivation to comply with the referent in a given behavioral domain. I think that the kind of research that Mike has been doing will be most helpful in resolving these questions, and I'd like to see a replication of his study where motivation to comply was again measured at different levels, but where the scale used was truly bipolar in nature.

The Relative Weights of the Two Components. The final set of problems relates to the meaning of the weights attached to the attitudinal and normative components as factors influencing intentions. Olli argues that since these weights are assumed to vary across individuals and behaviors, they are not parameters but variables, and thus multiple regression procedures for estimating the weights are inappropriate. Rich also questions the use of a regression approach as a means of estimating weights, but on somewhat different grounds. He points out that since SN and  $A_B$  are usually highly correlated, their multicollinearity raises serious questions about the reliability and validity of the obtained regression weights.

While there is some truth in what they say, I'm not quite sure what to do about it. One thing I do

know, however, is that I do not want to ask people to tell me how much importance they place on each component in determining their intentions. There is already more than enough data coming from studies of cue utilization to demonstrate that people cannot provide accurate estimates of the relative weights they place on different cues in arriving at judgments.

Further, I must admit I am not terribly concerned about the "multicollinearity problem". That is, I think in most cases in the "real world" there are "true" correlations between subjective norms and attitudes toward behaviors, and I think that any attempt to eliminate this correlation would be inappropriate. My current position is that I view the weights as theoretical parameters (parameters can vary), and correctly or incorrectly, I see multiple regression as one way of estimating these parameters. Despite the many problems associated with the multiple regression procedure, we have been able to show that weights obtained through multiple regression analyses do covary in accordance with our theoretical predictions. For example, consistent with expectations, SN was found to be more important (received a higher regression weight) than  $A_B$  in cooperative situations, while it was less important than  $A_B$  in competitive situations. Similarly as predicted, authoritarians did tend to place more weight on the normative component than did egalitarians. Given findings like this, I feel that despite some obvious inadequacies, the use of multiple-regression procedures to estimate the relative weights of the two components has yielded useful and meaningful results and in the absence of a viable alternative I will continue to use this procedure. Let me make it clear however, that I would really like to find some independent method for estimating or measuring the weights for each individual with respect to any given behavior. Unfortunately, I don't know how this can be done at the present time, but I do think this is a very important area for research.

Before concluding, there is one additional point I would like to make. In the course of the session, both Mike and Olli referred to my concern with specificity. Olli said that if I was concerned with specificity I should measure attitudes toward "my performance" of a behavior rather than "performing" the behavior per se, and Mike argued that to be consistent with my concern with specificity he feels that motivation to comply should be measured at the most specific level. Since I agree with Olli and disagree with Mike, it should be clear that my concern is not with specificity. However, many other people have also criticized the model for being too specific, and they have questioned the value of a "specific" concept such as an attitude toward an act. I would like to take this opportunity to point out that I am not now, nor have I ever been, concerned with specificity. What I am concerned with is correspondence. The model is a general model, and it is designed to account for people's intentions to perform any act. From a theoretical point of view, it makes no difference whether we are trying to predict a very specific intention (e.g., the intention to buy Crest toothpaste the next time I go shopping at my local A & P) or a very general intention (e.g., the intention to buy toothpaste). Once an intention has been identified, however, the model does state that all other measurements must correspond to that intention if one wants to obtain satisfactory prediction.

If I'm interested in a consumer's intentions to "buy toothpaste", then, according to the model, I must measure his or her attitude toward "buying toothpaste" and his or her subjective norm with respect to "buying toothpaste". Similarly, if I want to know the beliefs underlying this attitude or the subjective norm, I should be assessing his or her salient beliefs about "buying toothpaste" or his or her normative beliefs about "buying toothpaste". On the other hand, if I'm interested in a consumer's intention to "buy Crest toothpaste the next time I go shopping at my local A & P", the model says I must measure attitudes toward "buying Crest toothpaste the next time I go shopping at my local A & P" the subjective norm concerning "buying Crest toothpaste the next time I go shopping at my local A & P", etc. To put this somewhat differently, behavioral criteria and intentions can be described in terms of four elements: the action, the target the action is directed at, the situation or context in which the action occurs, and the time at which the action occurs. All four elements can vary along dimensions of specificity. For example, the target may be a specific brand (e.g., Wheaties) a relatively specific product class (e.g., dry cereals), a more general product class (breakfast cereals), a fairly global product class (food), or it can be left unspecified. The model doesn't care what level of specificity is adopted. However, once a behavioral criterion has been identified (i.e., defined in terms of the four elements), the intention must correspond directly to that criterion. Similarly, measures of beliefs, attitudes, and subjective norms must also correspond to the intention and the behavior. It is only under conditions of correspondence that maximal prediction can be expected. The actual level of specificity of the elements one deals with should be determined by the problem or behavior you are investigating.

#### References

- O. T. Ahtola, "The Vector Model of Preferences: An Alternative to the Fishbein Model," Journal of Marketing Research, 12, 1975, 52-59.
- I. Ajzen and M. Fishbein, "Attitudes and Normative Beliefs as Factors Influencing Behavioral Intentions," Journal of Personality and Social Psychology, 21, 1973, 1-9.
- I. Ajzen and M. Fishbein, "Attitude-behavior Relations: A Theoretical Analysis and Review of Empirical Research", unpublished manuscript, 1975.
- J. R. Bettman, N. Capon, and R. J. Lutz, "Information Processing in Attitude Formation and Change", Communication Research, 2, 1975, 267-278.
- A. Cowling, A. "Use of Elicitation Technique for Producing Dimensions of Brand Choice", Sixteenth annual conference papers, Market Research Society of Great Britain, 1973.
- M. Fishbein and I. Ajzen, Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research (Boston: Addison-Wesley, 1975)
- C. A. Insko, R. R. Blake, R. B. Cialdini, and S. A. Mulaik, "Attitude Toward Birth Control and Cognitive Consistency: Theoretical and Practical Implications of Survey Data", Journal of Personality and Social Psychology, 16, 1970, 228-237.

K. J. Kaplan and M. Fishbein, "The Source of Beliefs, Their Saliency and Prediction of Attitude", Journal of Social Psychology, 78, 1969, 63-74.

V. Kothandapani, "Validation of Feeling, Belief, and Intention to Act as Three Components of Attitude and Their Contribution to Prediction of Contraceptive Behavior," Journal of Personality and Social Psychology, 19, 1971, 321-333.

M. J. Rosenberg, "Cognitive Structure and Attitudinal Affect," Journal of Abnormal and Social Psychology, 53, 1956, 367-372.



## COMMENT ON THE MEANING OF "LIFE STYLE"

William D. Wells, Needham, Harper and Steers Advertising, Inc.

The term "life style" has been around for a long time, and has come to mean many things to many people. It was used extensively by Veblen, in his classic works on social class. To Veblen, life style was, in essence, what social class is all about.

The term "life style" has also figured prominently in the work of Alfred Adler. To Adler, life style was a much more individual concept than it was to Veblen. It meant something like "core personality" -- a way of responding to the world that the person develops early in life and that endures more or less unchanged over long periods of time.

More recently, "life style" has been widely and loosely used in the popular press. There it means expressive behavior -- usually public and usually involving the consumption of goods and services -- by which the individual defines to himself and to others his role in life. This use is similar to Veblen's but broader because it is not as focused on social class.

The term "life style" has also become attached to a type of research that attempts to provide detailed descriptions of consumers in terms of their activities, interests and opinions in addition to (not instead of) their demographic characteristics. The purpose of this research, some of which you will see later in this workshop, is to help communicators understand their audiences by providing detailed, multi-faceted descriptions that allow a marketer, a writer or a public policy maker to visualize the people he (or she) is trying to reach.

The activity, interest and opinion research, and the term "life style," developed separately. They came together because "life style" seemed to be such an appropriate shorthand expression for what the activity, interest and opinion research was attempting to portray. Thus, almost accidentally, the life style concept has become operationalized among a certain group of researchers as activity, interest and opinion research conducted for a rather limited set of purposes and employing a rather limited set of techniques. For other researchers, both within and outside of the realm of consumer behavior, the concept of life style would be operationalized in other ways.

There will no doubt continue to be groups of life style researchers whose aims and methods have little or nothing in common. This will distress the many of us who wish the world were simpler than it is; but, hopefully, it will benefit the field by maintaining competing points of view each of which has something to contribute that the others do not. At this point in its development, life style research cannot be tied into one neat operational package; that's good.

THE MEANING OF LIFE-STYLE:

SOCIOLOGICAL AND MARKETING PERSPECTIVES

Paul M. Hirsch, University of Chicago

Abstract

The concepts of image, status and prestige have long fascinated both marketing managers and academic sociologists. Historically, the term "life-style" has referred to a composite package of image, status and prestige elements which, for the sociologist, has signified a whole greater than the sum of its parts (Gusfield, 1963; Krugman, 1968). Viewed in this way, the concept has (appropriately) maintained widespread appeal at a theoretical level, and to the extent that the "whole" has presented difficulties in operationalization, we continue to use and accept it nonetheless for its intuitive appeal and face validity

The emergence of "life-style" studies in marketing is based primarily on a more empirical strategy of learning about the activities, interests and opinions (AIO) of consumers on a brand-by-brand or product-by-product basis, through structured questionnaires, often distinguishing heavy vs. light (or non-) users. The resultant profiles of "who" is typically the most frequent consumer of a specific brand or product thereby provides much information, though quite often on within-group variance among groups of respondents whom sociologists would likely characterize as sharing essentially similar rather than different life-styles. Whereas a sociologist, for example, might be interested in learning the AIO characteristics of users vs. non-users of all types of personal care products, he would be less concerned with the real marketing issue of what distinguishes brand-loyal or heavy users of Colgate from Gleem or UltraBrite toothpaste. This difference in outlook, however, while likely to provoke terminological disputes over the definition of life-styles, should not overshadow the present value and potential of AIO items and their utilization in consumer surveys as operational indicators in both marketing and "pure" research contexts.

One project at the University of Michigan Survey Research Center in which we (Robinson and Hirsch, 1975) found such life-style measures of substantial utility involved an effort to "predict" the distribution of marijuana use among teenagers. The SRC "Youth in Transition" Survey followed a national panel of 1620 young men from tenth grade through one year past high school graduation. Results from 1970 found 34 percent reporting marijuana use, with users distributed in such a way that virtually no demographic variables were significantly associated with the use of this product. Neither were grades in school, college attendance, region of residence, etc., although usage by friends was highly associated. Among the AIO items on the lengthy questionnaire were two "counterculture" measures, concerning attitude toward the Vietnam War and taste in popular music. Both were associated with use, but taste preference in popular music proved superior to any other predictor (except usage by friends) in the entire battery: those listing two or more "protest rock" records as favorites were twice as likely to report marijuana use as those listing

none or one. Additionally, as shown in the accompanying table, when the young men's attitude toward the war was crosstabulated with musical taste preference, the latter continued to hold up as the best statistical predictor of usage. These AIO-type items were thus useful indicators and important aids for constructing a profile of youthful marijuana users.

TABLE 1

Percent Using Marijuana as a Function of Music Preference and Attitudes Toward the Vietnam War (Year = 1970)

TOTAL RESPONSE (N=1544)	Number of Protest Rock Records Among the Three Favorites				No favorite records	
	Three	Two	One	None		
VIETNAM ATTITUDE						
Opposed (N=513)	55%	68%	63%	46%	36%	33%
Neutral (N=498)	36%	54%	43%	26%	24%	24%
Favor (N=513)	18%	32%	23%	14%	12%	20%

Further development of this type of life-style research will proceed as such "single product studies" are combined so that more composite profiles of the culture and life-styles of entire groups in society may be constructed, either by factor analysis directly from survey results (as in the Newspaper Advertising Bureau, 1973) or from a more demographically-grounded starting point (e.g., variations in the beliefs and behavior of different ethnic groups). For both products and attitudes these profiles will also need to be tracked over time. With replication of specific items, we will then find out whether the associations of 1970 or 1975 still hold in 1980 (Wells and Cosmas, 1975). Tracking life-style changes longitudinally and across groups could also provide a continuing source of information about the prospects for classes of products and their markets, and the subjective side of social change as well. This type of research, already under way at several universities and advertising agencies, is likely to prove useful to marketing managers and social scientists and should be carefully monitored by both.

References

Joseph Gusfield, Symbolic Crusade: Status Politics and the American Temperance Movement (Urbana: University of Illinois Press, 1963).

Herbert Krugman, "Consumer Behavior," in International Encyclopedia of the Social Sciences (New York: The Macmillan Company, 1968).

Newspaper Advertising Bureau, Psychographics: A Study of Personality, Life Style, and Consumption Patterns (New York: Newspaper Advertising Bureau, 1973).

John Robinson and Paul M. Hirsch, "Behavioral Correlates of Consumer Preferences in Popular Recordings: An Empirical Study of Diffusion Through the Mass Media" (mimeo: University of Chicago Graduate School of Business, 1975).

William Wells and Stephen Cosmas, "Life Styles: A Paper Prepared for the Consumer Behavior RANN Program" (mimeo: University of Illinois Survey Research Laboratory).

THE ADVANTAGES AND DISADVANTAGES OF THE  
PROFILE APPROACH TO ANALYZING LIFE STYLE DATA

Stephen C. Cosmas, Needham, Harper and Steers Advertising, Inc.

Abstract

This paper discusses the relative merits and demerits of using the profile technique to analyze life style data. The pros and cons discussed are those which are particularly relevant to profiling and not data analysis in general.

Today, I am going to talk to you about profiling, a way of analyzing consumers' life styles. Specifically, I want to discuss the advantages and disadvantages of the profiling technique. But, before I begin the discussion, it might be helpful to make some general comments about what profiles are and how they are constructed.

In a critical review article on "psychographics", Wells, (1975) made a comment to the effect that almost all of marketing is communication and that marketers are most effective when they know their audiences. Now, one technique for knowing an audience, and probably one of the most widely used for this purpose, is the life style profile.

The life style profile is a technique for analyzing life styles. To construct a profile, a target audience of interest must first be defined. The target audience can be the heavy users of a product or service, such as fast food restaurants (Tigert, Lathrope and Bleeq, 1971); readers of a particular magazine (Michaels, 1973); members of a specific age group (Wells and Cosmas, 1975); consumers who shop outside their own communities (Reynolds and Darden, 1972) -- in fact, any group of interest. Next, the target audience responses to the activity, interest and opinion questions are compared with the responses of the remainder of the population. Items which show significant differences are used to construct a profile of that target audience. Table 1 is one such example. This was a profile developed by Plummer (1971) of male bank charge card users.

Table 1

Cross-Tabulation Results of AIO Agreement  
With Male Bank Charge Card Usage

Statement	Card Users Definite and General Agreement	Noncard Users Definite and General Agreement
I enjoy going to concerts	25%	17%
A woman's place is in the home	27	41
In my job I tell people what to do	53	21
I am a good cook	36	26
My greatest achievements are still ahead of me	56	42
I buy many things with a charge credit card	39	22
We will probably move once in the next five years	46	37

Table 1 (cont'd)

Cross-Tabulation Results of AIO Agreement  
With Male Bank Charge Card Usage

Statement	Card Users Definite and General Agreement	Noncard Users Definite and General Agreement
Five years from now the family income will probably be a lot higher than it is now	71	60
Good grooming is a sign of self respect	52	71
There is too much advertising on TV today	59	70
Women wear too much make-up today	43	51
My job requires a lot of selling ability	51	37
I like to pay cash for everything I buy	26	67
Television is a primary source of our entertainment	25	40
Investing in the stock market is too risky for most families	47	56
To buy anything other than a house or car on credit is unwise	29	47
Young people have too many privileges today	52	64
I love the outdoors	54	76
There is too much emphasis on sex today	52	64
There are day people and there are night people; I am a day person	58	69
I expect to be a top executive in the next ten years	44	27
I am or have been president of a society or club	51	36
I would like to have my boss' job	42	33
A party wouldn't be a party without liquor	29	17
I would rather live in or near a big city than in or near a small town	46	34
I often bet money at the races	18	8
I like to think I'm a bit of a swinger	38	26

Table 1 (cont'd)

Cross-Tabulation Results of AIO Agreement  
With Male Bank Charge Card Usage

Statement	Card Users Definite and General Agreement	Noncard Users Definite and General Agreement
I stay home most evenings	62	71
Advertising can't sell me anything I don't want	55	68
I often have a cocktail before dinner	36	20
I like ballet	26	16
When I must choose between the two, I usually dress for fashion, not comfort	19	10
Liquor is a curse on American life	34	49
Movies should be censored	41	57
I read one or more business magazines regularly	34	18
I am active in two or more service organizations	28	17
I do more things socially than most of my friends	19	10
We often serve wine with dinner	30	16
I buy at least three suits a year	25	11
Playboy is one of my favorite magazines	25	16
I spend too much time talking on the telephone	31	17
It is good to have charge accounts	33	21
Hippies should be drafted	48	61
When I think of bad health, I think of doctor bills	31	46
My days seem to follow a definite routine	47	58

Here, the profile of male bank credit card users was produced by comparing that target group with the remainder of the population and taking those items that showed significant differences.

The profile shows, according to Plummer that, more than non-users, the heavy users of commercial bank credit cards agreed with items like:

"I would rather live in or near a big city than in or near a small town."

"I often have a cocktail before dinner."

"We often serve wine at dinner."

"I enjoy going to concerts."

"I like to think I'm a bit of a swinger."

"I expect to be a top executive in ten years."

"I do more things socially than most of my friends."

"I am or have been president of a club or society."

And disagree with items like:

"I stay home most evenings."

"There are day people and there are night people, I am a day person."

"My days seem to follow a definite routine such as eating meals at a regular time."

When added to the traditional demographics, this profile provides an interestingly detailed portrait of the male bank charge card user. The picture that emerges is one which seems:

"to typify the popular stereotypes of the successful man of the rise....

The picture of the suburban businessman arriving home from the office and having a cocktail, settling down to a nice meal, and then going off to various activities....

He is a busy, young businessman on the rise who knows where he is going....  
(Plummer, 1971)

In the above example, the independent variables (the life style questions) were related to the dependent variable (bank charge card use) through the simple technique of cross-tabulation. This represents one way of developing profiles. Other ways that have been used include correlational analysis, discriminant analysis, canonical correlation and combinations of the above (Tigert, 1966; Pessemier and Ginter, 1973; Darden and Reynolds, 1972).

However, rather than dwell on methods for constructing profiles, let me proceed to discuss the relative merits and demerits of using profiles, however constructed, to measure the life styles of target audiences.

One of the advantages of using profiles to measure consumer life styles is that once constructed, they are relatively easy to communicate to the users of the information.

Communication is relatively easy because life style profiles are interesting in and of themselves. They seem to describe how members of the target group live and interact with their environment. They describe peoples' activities, interests and opinions on a variety of everyday matters which are in the realm of experience of most individuals. This easy communication is an important advantage because if the user does not understand the profile, he is likely not to use it (Ellis, 1975).

A second advantage of profiles is that they focus on only one target -- the group that has been previously designated as the one of primary importance. This makes it possible to focus in on a single target rather than to fragment attention among several targets, as is the case with segmentation. Profiling, however, is not characterized by only advantages. There are several disadvantages to using profiles.

One disadvantage to using profiles is that they often lead both the researcher and the user to concentrate on differences and thus ignore absolute levels of response. For example, the item "I attend church

regularly," might produce the following hypothetical results:

	% Agree	
	Total Smokers	Non-filtered Smokers
I attend church regularly	74	63

From the response, one might conclude that non-filtered smokers are relatively uninterested in religion when compared to other smokers. As a comparison the conclusion is valid; but the temptation is to go even further here and conclude that non-filtered smokers do not go to church. In terms of absolute levels, that conclusion is obviously incorrect.

Looking back at the profile of male bank charge card users, it can be seen that when the absolute values of some items are re-examined, different conclusions may be reached. For example, the idea of the young businessman coming home from the office and having a cocktail might lead to the conclusion of a regular pattern of consumption of cocktails by the bank charge card users. But looking at the item "I often have a cocktail before dinner" it can be seen that only 36% agreed with the statement. This is not to single out this particular study for criticism, or even to say that the data were misinterpreted, but rather to point out that when profiles are used in the analysis of life style data, absolute levels can often be lost and their importance misplaced.

Probably the biggest disadvantage of profiles is that they may force the combination of several different groups into one. For example, among heavy users of soft drinks, there might exist several quite different groups of consumers with different life styles. If these groups are combined into one "heavy user" group, the differences between the groups are lost. Furthermore, if two different groups of heavy users differ from the non-users on the same dimension, but in opposite directions, merging the different heavy user groups into one overall heavy user group may lose the distinction entirely.

In summary, there are several advantages to using the profile approach to analyzing life style data. One, profiles are easy to communicate. And two, the profiles focus in on one target and therefore tend to present a more detailed picture of a target audience of interest. On the other hand, there are several disadvantages to using profiles. Researchers and users can be misled by concentrating on differences and ignoring the absolute value of the data. And, profiles may force the combination of several different life style groups and thus, sometimes cause the loss of important distinctions. One way of correcting this latter problem is through segmentation analysis, which will be discussed by Sunil Mehrotra.

#### REFERENCES

1. William R. Darden, and Fred D. Reynolds, "Canonical Correlation and Redundancy Analysis of Consumer Innovativeness and Life Style Characteristics," unpublished working paper, University of Georgia, 1972.
2. Richard Ellis, "Harnessing Pegasus: The Management of Multivariate Analysis," unpublished paper presented at the 58th International Marketing Conference of the American Marketing Association, Chicago, 1975.
3. Peter W. Michaels, "Life Style and Magazine Exposure," Combined Proceedings: Marketing Education and the Real World and Dynamic Marketing in a Changing World. Boris W. Becker and Helmut Becker, eds., Chicago, American Marketing Association, 1973, 324-331.
4. Edgar A. Pessemier, and James L. Ginter, "Profiles of Market Segments and Product Competitive Structures," Paper No. 409, Krannert Graduate School of Industrial Administration, Purdue University, May, 1973.
5. Joseph T. Plummer, "Life Style Patterns and Commercial Bank Credit Card Usage," Journal of Marketing, 35 (April, 1971), 35-41.
6. Fred D. Reynolds, and William R. Darden, "Intermarket Patronage: A Psychographic Study of Consumer Outshoppers," Journal of Marketing 36 (October, 1972), 50-54.
7. Douglas J. Tigert, "Psychometric Correlates of Opinion Leadership and Innovation," unpublished working paper, University of Chicago, 1969.
8. \_\_\_\_\_, Richard Lathrope, and Michael Bleeq, "The Fast Food Franchise: Psychographic and Demographic Segmentation Analysis," Journal of Retailing, 47 (Spring, 1971), 81-90.
9. William D. Wells, "Psychographics: A Critical Review," Journal of Marketing Research, 12 (May, 1975), 196-213.
10. \_\_\_\_\_, and Stephen C. Cosmas, "Life Styles," paper presented at the Conference of Knowledge of Consumer Behavior, Rann Program, National Science Foundation, April, 1975.

SEGMENTATION ANALYSIS -- A TOOL FOR MEASURING LIFE STYLES

Sunil Mehrotra, Needham, Harper and Steers Advertising, Inc.

Abstract

In this paper psychographic segmentation, as a tool for measuring life styles, is considered. A specific segmentation study done at Needham, Harper and Steers is cited; and the assets and liabilities of psychographic segmentation in general, and in relation to "profiling" is discussed.

The necessity for segmentation arises because of the limited capacity of human comprehension. Segmentation for humans serves the same function as organizing data into bytes does for a computer. It facilitates storage and retrieval of large amounts of data. Cognitive psychologists call this chunking.

We, in our daily lives, categorize information naturally. We classify people by sex, color, nationality and what have you; we think of cars as luxury, compact and subcompacts. A recent letter to Ann Landers illustrates an imaginative use of segmentation. A high school girl wrote saying that the boys in her class had rated all the girls from 1 to 5 on various criteria -- which we shall let you imagine. Her complaint was she had been rated 5 and she felt it was unjustified. (An example of misclassification, a problem I shall touch upon a little later).

This preoccupation with classifying women is best explained by James Thurber in his book "Is Sex Necessary" published in 1929. Here is what he had to say and I quote "...successfully to deal with a woman a man must know what type she is. There have been several methods of classification none of which I hold thoroughly satisfactory, neither the glandular categories -- the gonoids, thyroids, etc. -- nor the astrological -- Saigtarrious, Virgo, Leo and so on. One must be pretty expert to tell a good gonoid when he sees one.

...of much greater importance is a classification of females by actions. It comes out finally, the nature of women in what she does - her little bag of tricks as one might say."

Thurber, in his wisdom classified women into five groups. He categorized them as (1) clinging vine type, (2) "don't dear" type, (3) outdoors type, (4) button hole twister type and (5) the quiet type.

We at NH&S looked at her little bag of tricks and came up with a battery of Activity, Interest and Opinion measures. This formed the questionnaire that was sent to a national sample of 2,000 women. We used the responses to this questionnaire to segment the female data.

Here is what we did -- for each pair of women we computed a correlation coefficient indicating the similarity of their responses to the 250 AI&O questions. We applied principal components analysis with varmax rotation to this person by person correlation matrix. The factors obtained are regarded as hypothetical "woman types"; the factor loadings represent the correlation of each woman with each type. The women are then grouped by seeing which "type" each woman is most highly correlated with.

We ended up by classifying all the women into five groups. Each group was then cross-tabulated with the entire questionnaire and five profiles obtained. The five groups, on the basis of their central tendency, were identified as - 1) Cynthia, the chic city dweller; 2) Ursula, the urbane urbanite; 3) Mildred, the militant mother; 4) Cathy, the contented country girl and 5) Thelma, the old-fashioned traditionalist.

As an example of the descriptive value of segmentation by psychographics a summary of the life style profile of Ursula the Urbane Urbanite appears in Tables 1 and 2.

Table 1

DEMOGRAPHIC PROFILE OF URSULA THE URBANE URBANITE		
	URSULA (368)	ALL WOMEN (1797)
AGE	%	%
Under 25	8	12
25-34 years	24	24
35-44 years	26	20
45-54 years	25	21
55 years and over	17	24
HOUSEHOLD SIZE		
Two	31	37
Three	21	22
Four	25	22
Five	23	19
EMPLOYMENT STATUS		
Employed full time	30	28
Employed part time	17	14
Full time homemaker	41	45
EDUCATION		
Post graduate degree	7	2
Graduated college	25	12
Some college	38	30
Graduated high school	23	40
Less than high school	4	15
OCCUPATION OF HUSBAND		
Professional	25	13
Managerial/Administrative	26	18
Clerical/Sales	12	10

Table 1 (cont'd)

DEMOGRAPHIC PROFILE OF URSULA THE URBANE URBANITE		
	URSULA (368)	ALL WOMEN (1797)
OCCUPATION OF HUSBAND (CONT'D)	%	%
Craftsmen	12	18
Retired	8	13
Laborers/others	17	28
<u>TOTAL ANNUAL HOUSEHOLD INCOME</u>		
\$25,000 or more	20	10
\$20,000-\$24,999	17	11
\$15,000-\$19,999	16	14
\$10,000-\$14,999	25	29
\$8,000-\$9,999	7	10
\$4,000-\$7,999	4	8
Under \$4,000	2	4
<u>POPULATION DENSITY</u>		
Central City (all SMSA's - 50,000+)	31	29
Outside Central City (urban - 2,500+)	51	43
Rural (under 2,500)	18	28

Table 2

PSYCHOGRAPHIC PROFILE OF URSULA THE URBANE URBANITE		
	URSULA (368)	ALL WOMEN (1797)
	%	%
I would rather live in or near a big city than in or near a small town	45	37
I think the women's liberation movement is a good thing	64	51
I believe in a marriage where husband and wife share responsibilities	57	45
People tell me I am good looking	73	63
I like to feel attractive to members of the opposite sex	91	84
I have more self-confidence than most of my friends	74	64
My greatest achievements are still ahead of me	69	63
Our family income is high enough to satisfy our important desires	82	72
I would like to take a trip around the world	90	67
Most of my friends have graduated from college	74	42
I am much happier now than I ever was before	84	78
Most people are honest	81	73
Advertising insults my intelligence	72	61
I am in very good physical condition	83	76
Magazines are more interesting	65	44

## Psychographic Segmentation -- Assets and Liabilities

I shall now focus on some of the problems encountered in segmenting data in this manner. I may add these problems are inherent in most clustering algorithms.

The major problem is one of reproducibility; that is, do the segments really exist or are they merely artifacts of the method of analysis?

We have handled this problem by doing a split half analysis. The data were randomly split into two halves, and the clustering algorithm run separately on the two halves. When the solutions obtained for the two halves were compared, it was found that solutions with more than five groups did not reproduce.

However, this still did not guarantee reproducibility with a totally different set of AI&O measures or a totally different sample.

Of course, we had the reassuring knowledge that our five group solution was replicated by the high school kids and James Thurber.

A related problem is the problem of groups sizes. Are the sizes unique or would a different clustering algorithm or different data set produce totally different numbers per group?

To get a handle on this problem we looked at the Chi-square statistics for the distribution of group sizes in the two halves. We found the group sizes in two halves were not significantly dissimilar.

Lastly, we have the problem of interpretation. Are the group central tendencies representative of the group? Are the groups homogenous enough to infer that most people in the group look like the group average? This is unlikely. Given attitudinal data, which are fuzzy, there are likely to be more people at the fringes of each group than around the average.

Despite its problems, we feel segmentation is a worthwhile way of looking at data.

It reduces considerably the dimensionality of the data space. This is important both for the researcher and the user of research, for it helps bring out the information in the data in all its richness without inflicting information overload.

Another advantage of segmentation is that it is objective. We let the data speak for itself. Therefore, objective instruments can be designed to validate the findings of segmentation.

Once the segments have been identified it provides a means for keeping track of consumer markets through trend analysis. One can determine which segments are growing and which shrinking. This would help in repositioning existing products and in identifying new opportunities.

Finally, segmentation gets around the problem often encountered in "profiling"; that is, several disparate groups get mishmashed into one profile and the profile emerges looking something like a film that has had multiple exposures.



## AN EXPLORATION OF THE ROLE OF FAMILY LIFE STYLE ON SELECTED BEHAVIOR VARIABLES

Lucy Chao Lee  
University of Illinois

Social scientists seem to be fascinated by the idea that a person's life style influences his various socioeconomic behaviors, however complex and subtle is the process of such influence. It is generally recognized that the most intriguing tasks for implementing the idea are: first, to develop a theoretical construct for the concept of life style which can be measured by real data, and second, to test empirically the meaningfulness of the measurement in terms of its effects on behavioral variables.

I would like to share with you the experience that Professor Ferber and I had in measuring, and testing the effect of, family life style of young married couples in two areas of Illinois. The objective of this ongoing study is to measure the dimensions of family life style from an operational point of view, and then to test the significance of these dimensions used alone and in conjunction with other demographic variables, such as stock or purchases of durable goods, financial assets, etc.

The conceptual framework used in this study is to formulate the concept of family life style in very general terms. It is felt intuitively that family life style is, on the one hand, influenced by socioeconomic characteristics of the family and restricted by its given financial resources, and on the other shaped by the goals and attitudes of the members of the family. However, without looking into the multitude of the determinants of family life style, three general dimensions are postulated, as follows:

1. Career orientation
2. Pleasure orientation
3. Home orientation

These three dimensions by no means preclude each other. While a conglomeration of all three may be an axis about which family life revolves, each individual dimension will exert its influence on certain types of behavior of the family as a unit.

The available data are from two panels of young married couples. One panel comprises couples married in the summer of 1968 in the cities of Decatur and Peoria, Illinois, and the other, a younger panel, comprises couples married in the summer of 1972, in the Chicago area; in all cases with the husband 30 years of age or less at the time of marriage. These couples have been interviewed every six or twelve months since their marriage, and a great deal of information has been collected.

In particular, uses of time as reported in the Spring of 1973 by the husband and wife in 11 activities served as the basis for measuring the dimensions of family life style. Table 1 shows that the Decatur-Peoria couples were spending substantially more hours per week on housework and children than the Chicago couples. This is expected, because these couples were married 4 years earlier, and most of them already had children and were homeowners. On the other hand, the Chicago couples were spending relatively more time on job, travel, and visit and entertain. Even so, to some extent we can use one panel as a validation sample for checking the statistical results obtained from the other panel.

To condense the 11 use-of-time variables, factor analysis was applied to their correlation matrix. Three common factors were extracted by the principal axes method, and then rotated orthogonally by the varimax method for better interpretation. Table 2 shows the varimax solution for each panel. Listed under each factor are use-of-time variables in descending order with absolute loadings greater than 0.4; variables with negative loadings are in parentheses. The results indicate that the constituents of the three factors are very similar between the two panels. They seem to fall in line with the 3 postulated dimensions, and are therefore labeled as such.

The next step was to obtain the regression estimates of the factor scores, which were then used as the basis for classifying the families into types or segments. Table 3 lists three criteria for classification that were developed for this purpose. The first criterion identifies each family by the single factor with the highest score. For example, a family which scores highest on career orientation is classified as leading Type (1) life style, labeled by C, career oriented. Naturally, only three types or segments resulted from this criterion.

The second criterion identifies each family by the order of the three factor scores according to their magnitudes. For example, a family classified by this criterion as leading Type (1) life style labeled by CPH means that the family has the highest factor score on career orientation, next highest score on pleasure orientation, and lowest score on home orientation. Six types or segments resulted from such a permutation.

The third criterion differentiates among the sign combinations of the three factor scores. For example, families with all three positive or zero factor scores are classified as leading Type (1) life style, labeled by +++. By this criterion, eight types or segments are obtained.

The meaningfulness of these three classifications of family life style, and of the factor scores themselves, was tested by using them as four sets of independent variables in separate regression functions for explaining certain behavioral, dependent variables. The dependent variables tested for the Decatur-Peoria panel were stock and purchases of durable goods, financial assets, satisfaction in life and number of children. Only two statistically significant functions were obtained (at the .10 level). Both relate to the third criterion of life style, and they are shown in Table 4. One is number of durables owned, significant at the 10% level, and the other is total debts of the family, significant at the 5% level. Note the significant variables in each function. The beta coefficient of life style Type (6) is positive in the first function, which implies that families with more durables spend relatively more time on pleasure activities but less time on career and home activities. In the second function, life style Type (3) has a negative beta coefficient, which means that families with less debts spend relatively less time on pleasure activities but more time on career and home activities.

The results obtained so far are not spectacular, and yet not unreasonable. Much work has yet to be done.

Other behavioral variables will be used, such as exposure to media, vacation plans, impact of energy crisis of life, etc. We also intend to compare the explanatory power of life style variables with that of the usual demographic variables.

TABLE 1  
COUPLE'S COMBINED USE OF TIME (AVERAGE NUMBER OF HOURS PER WEEK), BY PANEL AND ACTIVITY

	Chicago	Decatur-Peoria
1. Housework	19.8	34.9
2. Shop for home needs	8.6	7.4
3. Job	63.6	53.9
4. School	2.2	0.3
5. Travel for job or school	8.3	4.4
6. TV and hifi	23.9	24.3
7. Indoor recreation	12.2	11.0
8. Outdoor spectator sports	5.2	3.7
9. Outdoor participant sports	4.9	4.7
10. Visit or entertain	16.5	11.9
11. Childcare	20.9	69.3
Sample Size	184	134

TABLE 2  
VARIMAX SOLUTION FOR 11 USE-OF-TIME VARIABLES

Common Factors	Constituent variables	
	Chicago	Decatur-Peoria
Career orientation	3. Job	3. Job
	5. Travel	1. (Housework) 5. Travel
Pleasure orientation	7. Indoor recreation	10. Visit or entertain
	6. TV or hifi	9. Outdoor participant sports
	8. Outdoor spectator sports	6. TV or hifi 7. Outdoor spectator sports
Home orientation	11. Childcare	11. Childcare
	1. Housework	7. Indoor recreation
	2. Shopping	

TABLE 3  
THREE CLASSIFICATIONS OF LIFE STYLES  
BASED ON REGRESSION ESTIMATES OF FACTOR SCORES

1. Using the factor with the highest score	2. Listing 3 factor scores in descending order	3. Noting the signs of 3 factor scores
		<u>C P H</u>
(1) C	(1) CPH	(1) + + +
(2) P	(2) CHP	(2) + + -
(3) H	(3) PCH	(3) + - +
	(4) PHC	(4) - + +
	(5) HCP	(5) + - -
	(6) HPC	(6) - + -
		(7) - - +
		(8) - - -

C = Career orientation  
P = Pleasure orientation  
H = Home orientation

TABLE 4  
BETA COEFFICIENTS OF REGRESSION FUNCTIONS,  
DECATUR-PEORIA PANEL

Type of life style	Dependent variable	
	No. of durables	Total debts
<u>C P H</u>		
(1) + + +	0.14	-0.12
(2) + + -	0.00	0.14
(3) + - +	-0.16	-0.28*
(4) - + +	-0.04	-0.19
(5) + - -	0.13	0.01
(6) - + -	0.23 <sup>Δ</sup>	-0.08
(7) - - +	0.15	0.00
Adjusted R <sup>2</sup>	0.05 <sup>Δ</sup>	0.08*

<sup>Δ</sup> Significant at 0.10 level.  
\*Significant at 0.05 level.  
#Omitted is life style type (8), - - -

CHILDREN AS CONSUMERS: THE NEED  
FOR MULTITHEORETICAL PERSPECTIVES<sup>1</sup>

Thomas S. Robertson, Center for Research on Media  
and Children, University of Pennsylvania  
Shel Feldman, Brooklyn College

Abstract

Our objective in designing this workshop has been to confront the issue of how to encourage interdisciplinary research on children's consumer behavior. The particular task of this paper is to posit some direction for resolving the conflict and inconsistency engendered by the specification of multiple theoretical perspectives. Our approach will be to suggest that not all components of consumer behavior, nor all aspects of a social issue are equally amenable to a given theoretical perspective. Instead, a critical function for the researcher is to match particular problems with appropriate theoretical vantage points -- a function which many researchers abdicate due to their faithful allegiance to specific theoretical positions.

Children as Consumers

The topic of children's consumer behavior is one of some currency in social policy circles and has sparked a flurry of research in the past five years. Yet, much of the data which has been generated has been free of any theoretical underpinnings -- which is perhaps to be expected in an embryonic research tradition. Our position is that social policy will be served best in the long run by research which is theoretically-based, rather than in specific response to issues of the moment.

Children's consumer behavior is a topic which is rich in theoretical and methodological implications. The sheer pervasiveness of consumption and the range of possible behaviors suggests a rich laboratory and research base for the application of many diverse theoretical vantage points. Furthermore, consumer behavior may be pervasive not only socially, but in many cases it may be an important mode of self-presentation and self-enhancement. The corresponding complexity of consumer behavior also argues as to its potential theoretical robustness. The dynamics of consumer decisions may involve us in alternative and sometimes conflicting theoretical perspectives, since no single vantage point is capable of broad levels of prediction.

Theoretical Allegiance

The researcher often finds that progress in understanding a problem results from taking a particular perspective on that problem and extending it to its limits. Such an exercise clarifies the assumptions involved; elucidates the problems to which the perspective does not apply; and organizes the data on the problems to which it does apply. This procedure works in part because of its selective effect. Theoretical commitment limits the phenomena accepted for study, the methods used, and the classes of explanation

found worthy of consideration.

A theory often embodies a researcher's perspective on the world: given one perspective, he investigates certain phenomena; given another perspective, he investigates quite different phenomena. The differences between the phenomena studied by different researchers often preclude definitive objective tests of the perspectives themselves. For this reason, among others, theories seem more often to fall as victims of revolution than to be indicted by normal processes of evidentiary examination.

The role taken by the researcher in publishing, and in the usual interdisciplinary conference (such as this), is as explainer of, apologist for, and crusader on behalf of, his own theoretical perspective. Even neglecting the researcher's ordinary tendencies to defend his own viewpoint in contention with others, it must be noted that he is trained to approach problems from a single theoretical perspective. If the several participants in such a conference each take such roles, it is not surprising to find that attempts to integrate their contributions into a coherent whole are generally unsuccessful. It must be remembered, however, that we are not discussing pathologically closed-minded persons, or even zealots with content commitments, but those who have found a particular intellectual vantage point highly functional. In the clash of perspectives with his colleagues, the scientist finds the apparent limits to his own perspective and is challenged to extend that perspective to overcome them. He feels the possibilities for growth and development and begins to meet some of these -- and herein lies the individual stimulation and learning.

Functional as such practices may be for the individual researcher, or for the development of pure knowledge, they may be dysfunctional for the solution of social problems, or marketing management problems. Real phenomena do not come neatly encapsulated, but as a congeries of effects on many facets of social, cultural, economic, and political norms, attitudes, and behavior. Given his training, the scientist attempts to deal with the multiplicity of causes and effects first by separating out those aspects of the problem that are most amenable to exploration within his own perspective, and then by either analogizing to the remainder or treating it as residual variance. Such an approach works well, it must be reiterated, in advancing the understanding of a particular perspective, or the understanding of the individual researcher.

Piagetian Theory: A Case Example

Let us consider, as a straw man example, the Piagetian perspective on development. The study of children as consumers seems a natural candidate for extension of Piagetian theory. Much of the child's behavior with respect to consumption must depend upon his knowledge about money and exchange relationships; his ability to compare products on various attributes and evaluative

<sup>1</sup>The authors are indebted to the Consumer Research Institute for funding.

dimensions; and his ability to plan and balance requirements for consumption and for saving. The perspectives on the development and functioning of the cognitive structures underlying these achievements offered by Piaget and his followers are surely pertinent to the issues raised. These perspectives are likely to cast considerable light on the abilities the child may bring to bear on his problems in the sphere of consumption.

At the same time, Piaget and his followers have been little concerned about the affective life of the child and this perspective therefore would not seem to be particularly useful in examining advertising's impact on the child's desires, fears, and anxieties. So long as Piagetian theory's major concern is with respect to accommodation to the physical world, little is lost by this choice among phenomena. Even when concern shifts to accommodation to modes of moral reasoning, little need be lost, insofar as such reasoning is universal (which is not, of course, to say that the contents of the values involved are universal) and there is little conflict between models available to the child. But the essence of the problem here is that alternative models of consumer behavior and reasoning about consumer behavior seem to be made available to the child, presenting occasions for considerable conflict and affective imbalance.

The incompleteness of the Piagetian framework is further argued by Calder, Robertson and Rossiter in a paper following this one. They argue that Piagetian theory denies the mediational representational view of information processing, whereby information from the external environment is "represented" psychologically. Piaget does not subscribe to the thesis that language is the basis of thought nor that cognitive structures depend on symbolic representations. It may be that Piaget's biological approach and the representational view are not incompatible, but they are definitely appropriate for different problems. What one cannot do is to impose cognitive development stages on the typical consumer information processing model. The child's current cognitive responses may be regulated by structures, or internalized operational schemes, but they are not explained by structural changes. Stage theory is important for understanding long-term developmental changes, not necessarily for understanding short-term shifts.

Is the sort of insufficiency just discussed peculiar to Piagetian theory? Certainly not! Any theoretical perspective is partial with respect to its explanatory power and the range of phenomena to which it can be applied. Children's consumer behavior cannot be handled exclusively within the Piagetian framework any more than it can be handled exclusively within a socialization, social learning, psycholinguistics, or whatever type of theoretical perspective.

#### Multitheoretical Integration

How are we to choose among theoretical perspectives or to find their underlying commonalities? The most popular current models of consumer behavior suggest integration by the application of some meta-theory or systems theory.

It is our suggestion that the generation of understanding about children's consumer behavior requires a modification of the systems analysis approach. In this approach the analyst attempts not to isolate component processes of the entire system, but rather those questions appropriate to each of the theoretical perspectives offered. In our view, a normal systems approach requires an a priori theory -- or meta-theory

-- of the entire problem, and this is generally not available for the class of behavior which we are considering, and would often be inappropriate if it were available, for exactly the same reasons as given above for the insufficiency of the usual perspectives brought to bear on consumer behavior problems.

Our thesis is that all aspects of consumer behavior are not equally amenable to a given theoretical perspective. The researcher must sort among the perspectives available and choose those which best fit the behavior he is most interested in explaining. In doing so, he must also recognize that the behavior is not fully explained, but only some portion or portions thereof. It follows that, as his resources grow, he can research several aspects of the problem, and that the perspectives applied to each of these need not be the same.

It is our thesis, therefore, that children's consumer behavior must be divided into its component parts in such a way that problems are matched with theoretical perspectives. Let us consider the problem of children's consumer behavior more directly. What are some of the theoretical perspectives available, and what are some of the component problems that fit each perspective best? We shall attempt a brief illustrative review.

This review is organized by theoretical perspectives on children's consumer behavior from the behavioral sciences. It is not an attempt at a rigid categorization of these perspectives: theories cannot be categorized along a single dimension; neither are their concerns exhaustive nor mutually exclusive. This review is meant only to be suggestive of the sorts of issues with which particular broad theories seem most concerned. It covers those we shall loosely label as: learning theory, cognitive-development theory, information-processing theory, psychoanalytic theory, attitude theory, socialization theory, and group process theory. It is assumed that the reader knows or has access to full descriptions of these theories; what is attempted here is to sketch the general concerns of each, and to present a number of specific questions about children's consumer behavior that seem congruent with each.

#### Learning Theory Approaches

The major concerns of learning theorists relate to the growth of specific skills and their use in particular situations. Some simple illustrations of this concern would be the study of children's learning of brand names, product uses, and commercial message content and slogans. More interesting illustrations are studies of the ways in which the child first learns the connection between commercial messages and the purchase of consumer goods; the connection between the content of such messages and the experience of using the advertised products; and the lack of connection between such messages and the content of the programs in which they are embedded. The learning theorist is also concerned with such problems as determining how the child comes to discriminate among products of similar function and to generalize brand images across products. Problems of even larger scope might also be studied, such as the child's learning to save his discretionary funds and the strength of this response as opposed to the consumption response.

Particularly relevant is social learning theory involving stimuli provided by people. Atkin, in a paper which follows, distinguishes between instrumental training where a "teacher" explicitly attempts to shape response by differential reinforcement and imitation, where the child matches responses to cues provided by responses of a model. A considerable proportion of the

child's learning comes about based on imitation.

The position of parents in the social learning process is both as teacher and role model. It becomes important, therefore, to study the child's learning of responses to parental reactions regarding purchase requests; the gradual shaping of those responses into standard strategies; and the use of particular strategies in response to particular sorts of parental reactions. It is important to research what responses parents attempt to teach and reinforce. For example, parents may deal explicitly with their objections to requests for particular products, or attempt to teach the child the criteria they themselves are using, or they may allow the child to learn by trial and error. Parents may reinforce saving behavior or consumption, comparison shopping or impulse buying, and so forth, by their example, or by explicit rewards. They may use material rewards or symbolic rewards in their child-training practices, immediate rewards or deferred rewards. The responses suggested, the explicitness of the teaching, and the types of reinforcement offered for acceptable behavior may all be productively investigated from the perspective of social learning theory.

Media models as sources for children's learning also comprise an interesting topic of investigation. Atkin, for example, in his paper to follow, notes the role of adults in children's commercials as authority figures, who either directly approve of the child model's consumption, or implicitly endorse it by their presence. Such reinforcement has a bearing on the impact of the message. Research could also be conducted on source characteristics affecting message reception and yielding, such as the use of child actors as models or the relative effects of race or sex. Research to this point on such source characteristics is indeed sparse.

#### Cognitive-Development Approaches

As suggested earlier, the principal concerns of those who study cognitive development are to specify the competence the child brings to his behavior as a consumer, and to specify the course and determinants for the development of that competence. The developmentalist would be concerned, for example, with the child's understanding of money as a medium of exchange. Thus, while he would inquire, as would the learning theorist, as to how and when the child learns to identify coins as such, and to identify different sorts of coins, he would also inquire as to how and when the child learns to view coins as a storehouse of value and to appreciate their use in purchasing as meta-barter functioning (Strauss, 1952).

Another area of interest relates to evaluating products in a class and computing tradeoff functions among them. The child must learn about the intersubstitutability of various items. He must also be able to evaluate their merits on each of several attributes or dimensions, and to bear this multiplex in mind while choosing to purchase one or another of the items. If the child is to evidence effective consumer behavior, he must be capable of such complex cognitive functioning as these formalizations suggest.

Furthermore, it is important to the developmentalist to study the experiences that are necessary or sufficient for development of the abilities just discussed. He would also be interested in the sorts of behavior that occur prior to the development of those abilities, and in the fit of those other behaviors to environmental and interpersonal demands in different socio-economic circumstances and settings. The article by Ward, Wackman and Wartella which follows elaborates this

cognitive development perspective.

#### Information Processing Approaches

Cognitive theories of information-processing are highly prevalent in psychology, communications, and consumer behavior research today. The essential focus of such models is on how individuals select, store, evaluate, and utilize information for decision-making purposes (Mc Guire, 1969). The question arises as to how effective and "rational" an information-processing machine the child is in absorbing information, processing it, and modifying his behavior accordingly. Of concern here is the child's ability to procure information, his potential overreliance on limited sources (advertising), and his lack of cognitive defenses in sorting-out and evaluating sales-oriented information.

The paper by Ward, Wackman and Wartella which follows provides an overview of this perspective and provides a synopsis of children's information processing behavior. Also relevant is the paper by Fowles which follows in that it addresses the information processing consequences of television viewing as a potentially passive form of behavior. Given the child's lack of selectivity in processing televised information, since his attention is dictated by the stimulus structure. Fowles' raises the question of whether the message may be passively accepted: "If the sequence of percepts is largely dictated, then it follows that evaluation of these percepts (for plausibility, truth, value, etc.) is largely suspended, since judgment is not a passive mode of response."

Calder, Robertson and Rossiter (in another paper to follow) contrast adult versus child models of information processing. Their argument is that the child model is not simply a constrained version of the adult model but is both quantitatively and qualitatively different from that of adults. In particular, they argue that research on children's consumer information processing should not follow the pattern of adult research in ignoring all but the linguistic coding of information. Instead, considerable emphasis must also be placed on enactive (motor) and imagery codes. Rossiter, for example, in his paper to follow, finds evidence that children have a rich data base of information about cereal brands which is stored in visual rather than verbal memory and which impacts on preferences for brands.

#### Psychoanalytic Approaches

The central concerns of theorists who take the psychoanalytic approach are the affective and motivational dynamics of behavior. One problem of interest is the child's motives for consumption of particular products. Research might, for example, investigate whether the gratification which a child experiences in playing with a particular toy arises from its connection with previously cast-off infantile behavior, from the opportunities it offers to master the immediate environment, or from the opportunity it offers to play the adult role. Psychoanalytically-oriented researchers might study the child's use of, and choice among, consumer products as a means of allaying anxiety about sibling relationships or parental relationships.

Psychoanalytically-oriented theorists would be interested in the effect of commercial messages upon the child's fantasy life, including the degree of competitiveness, aggression, or dependency they fostered. They would be concerned about the manner in which family conflicts over purchase decisions fit into other ongoing conflicts and aid or retard the development of stable and effective modes of conflict resolution.

### Attitude Theory Approaches

One primary focus of attitude theory is on the prediction of behavior based on knowledge of an individual's attitudes -- generally conceptualized in terms of cognitive, affective, and action tendency components. Of particular interest in relating attitude theory to consumer decisions might be the balance theory approaches stemming from Heider and variously conceptualized in the communications' and consumer behavior literature. Also of interest is the recent literature on attribution theory (for example, Jones et al, 1971) which views intent from the vantage point of the perceiver and which can be extended to media communicator intent (Robertson and Rossiter, 1974).

An initial question is when and how children form attitudes toward product categories and how well these attitudes predict behavior. Furthermore, what are the relative influences in the formation of attitudes and what role does mass media advertising play? Even more specifically, how are brand preferences formed among children and again what role does advertising play? It might be hypothesized that mass media advertising serves a critical function in the creation of children's attitudes and preferences, especially if parents treat consumption as a low salience topic and are not actively involved in the process of inputting to and critiquing the child's developing attitude structures.

It has been suggested that children may be particularly subject to disappointment as a result of consumption (Feldman and Wolf, 1974). This is posited due to the belief that unrealistic expectations are built by advertising. The levels of disappointment incurred by children as a result of consumption and the dissonance-coping modes used constitute an interesting and researchable topic.

### Socialization Theory Approaches

Socialization refers to the process by which children learn a particular role and the requisite values and requirements for performance of that role. Beuf, in a paper to follow, suggests that the extensive use of television by American children "has the potential to function as a common conscience which introduces and reinforces American values at a high level of generality." Beuf's view is that television and advertising may be considered a new secondary socializer "stepping in to perform some of the tasks of value definition and inculcation which are no longer so effectively performed by sacred institutions."

Little is known in a rigorous sense about television advertising's longitudinal effects on values or its relative effects compared to other agents of socialization. Nevertheless, concern is voiced by social critics as to possible negative long-term socialization consequences. Some of the undesirable effects alleged are: that advertising instills undesirable values; that advertising encourages materialism and the need to consume; and that advertising does not teach rational consumer decision-making. Perhaps the most reasonable response by marketers to these charges has been that advertising only reflects the existing values of the society and does not create these values. Furthermore, advertising performs a vital function in preparing the child for his role as a consumer, at least in U.S. society.

Mass media are not an independent means of socialization and would generally be expected to be of less importance than family and peers -- although this depends on the age and circumstances of the child. The

young child who is deprived of a meaningful family context and who has been exposed to television as surrogate guardian and babysitter may be unduly influenced by this medium, form stereotyped impressions, and learn a limited value system regarding consumption (Gerbner and Gross, 1973).

### Group Process Theory Approaches

Ultimately, the child's consumption behavior is dependent on the dynamics of his family group. This suggests the possible usefulness of group process theories. Marketing to children is based primarily on encouraging the child to initiate a purchase request to his parents. According to some social critics, this may lead to an unhealthy family relationship -- particularly in low income homes where the parents must deny a disproportionate number of requests.

Exactly what happens in consumption-related interaction between parent and child is not well researched. The particular outcome of this interaction and the role and relative influence of television advertising is probably a function of many factors, including social class, the child's position in the sibling order, the age of the child, the degree of role integration between parents, and the parental discipline style. For example, general research on intra-family interaction processes finds that middle-class parents rely more on intellect and reasoning than on authority position and are, therefore, more open to interaction and suggestion from their children (Berkowitz, 1964). Preliminary research evidence by Ward and Wackman (1972) reveals that purchase requests to parents decrease as the child grows older but that the parent's yielding to the child's requests increases as the child grows older. Conflict level between parent and child rises with the number of requests which the child makes.

### Conclusion

It is tempting to apply behavioral science theory intact to the area of children's consumer behavior. Following Lewin's dictum, "Nothing is so practical as a good theory," the researcher attempts to cast all aspects of behavior into terms covered by his own theoretical approach. All too often, however, that which is not encompassed is then passed off as irrelevant, and the theory is thus confirmed by a selection from among the available data. Although theoretical perspectives do overlap, there is generally a fairly specified set of behavior to which any one theory applies.

In order to understand and explain the dynamics of children's behavior as consumers, it is necessary first to bring to bear an enriched set of theoretical perspectives and then to select those theories most appropriate to particular kinds of behavior. But this sort of selection cannot be done from within the confines of a particular perspective. As Godel has proven, no system at least as complex as elementary arithmetic is susceptible of being shown logically consistent, except from the vantage point of a more powerful system than itself. The theorist is unable to assess the ultimate consistency of his perspective purely within that perspective. Neither is he able to assess its consistency from simple comparison with an alternative perspective that makes quite different assumptions than his own. It is only as he observes his system in comparison with others and with certain simplifying constraints relaxed, as in considering its appropriateness to an applied problem, that its consistency can be assessed and, if necessary, modified.

In conclusion, we do not wish to destroy scientific singlemindedness, but instead to harness its potential productivity toward understanding children as consumers. To this end we have offered the suggestion that children's consumer behavior be divided into its component problems and that these problems be matched with appropriate theoretical vantage points.

#### References

Leonard Berkowitz, The Development of Motives and Values in the Child (New York: Basic Books, 1964).

Shel Feldman and Abraham Wolf, "What's Wrong With Children's Commercials?," Journal of Advertising Research, 14 (February, 1974), 39-43.

George Gerbner and Larry Gross, "Cultural Indicators: The Social Reality of Television Drama." A working paper, Annenberg School, University of Pennsylvania, 1973.

Edward Jones et al. Attribution: Perceiving the Causes of Behavior (New Jersey: General Learning Press, 1971.)

Mary G. Jones, "The FTC's Need for Social Science Research," Proceedings: Second Annual Conference, Association for Consumer Research, (College Park, Maryland: ACR, 1971).

William J. McGuire, "An Information-Processing Model of Advertising Effectiveness," A paper presented at the Symposium on Behavioral and Management Science in Marketing, University of Chicago, July 1969.

Thomas S. Robertson and John R. Rossiter, "Children and Commercial Persuasion: An Attribution Theory Analysis," Journal of Consumer Research, 1 (June, 1974), 13-20.

Abraham Strauss, "The Development and Transformation of Monetary Meanings in the Child," American Sociological Review, 17 (June, 1952), 275-286.

Scott Ward and Daniel Wackman, "Children's Purchase Influence Attempts and Parental Yielding," Journal of Marketing Research, 9 (August, 1972), 316-319.

CHILDREN'S SOCIAL LEARNING FROM TELEVISION ADVERTISING:  
RESEARCH EVIDENCE ON OBSERVATIONAL MODELING OF PRODUCT CONSUMPTION

Charles K. Atkin, Michigan State University

Abstract

This paper assesses advertising effects on children and adolescents from a social learning theory perspective, emphasizing imitative performance of vicariously reinforced consumption stimuli. The basic elements of Albert Bandura's modeling theory are outlined, and specific derivations from the theory are applied to the problem of television advertising effects; new research evidence is presented and interpreted in terms of social learning principles.

Social Learning Theory

Gewirtz (1969) defines social learning as a category of learning involving stimuli provided by people. There are two basic social learning processes that influence behavior: direct instrumental training, where a teacher explicitly attempts to shape responses via differential reinforcement, and imitation, where an observer matches responses to discriminative cues provided by responses of a model. The latter process occurs without direct tuition and comprises the vast majority of a child's socialization.

Although traditional learning theories portray behavior as a function of experiential reinforcements, Bandura contends that "virtually all learning phenomena resulting from direct experiences can occur on a vicarious basis through observation of other people's behavior and its consequences for them." This enables man to acquire behavior by example without tedious trial and error practice; similarly, behavioral inhibitions can be induced by observing punishment of others for their actions. There is an important cognitive component in this approach, as people symbolically represent and process external influences for later guidance in performance (Bandura, 1962, 1965, 1969, 1971a, 1971b, 1973; Bandura and Walters, 1963; Bandura, Ross, and Ross, 1963a, 1963b).

In Bandura's view, "most of the behaviors that people display are learned, either deliberately or inadvertently, through the influence of example." The reasons for this are that dangerous mistakes can be avoided by observing competent models who demonstrate proper modes of response, that complex behaviors such as language can only be learned through example, and that novel response patterns can be acquired more efficiently through observational means.

Early social learning and operant conditioning theories of modeling emphasized the requirement of overt performance of a response matching the modeled stimulus cue, followed by the positive reinforcement of the imitative behavior (Miller and Dollard, 1941; Skinner, 1953). Bandura's approach differs from contemporary analyses in the treatment of response integration, cognitive control, and reinforcement influences. He proposes four sub-processes that govern modeling. A distinction is drawn between acquisition (learning) and actual performance of the behavior. An observed response becomes acquired when stimuli merely elicit mental representations of the behavior; overt performance is primarily determined by reinforcements experienced by the model and the observer.

The initial process is attention, where the observer is exposed to the modeling stimuli. Acquisition of matching behavior at the sensory registration level occurs when the observer attends to, recognizes, and differentiates the distinctive features of the model's behavior. Attention is selectively determined by several sets of factors, including opportunity for encounter with direct or mediated models, the relevance, competence, distinctiveness, power, and attractiveness of these models, and the observer's sensory capacities, arousal level, motivations, and reinforcement history. Bandura feels that televised models are so intrinsically interesting that they are highly effective in attracting attention.

Retentional processes are crucial because delayed imitation outside of the model's presence requires symbolic representation in memory. This mental representation process involves coding of both images and verbal symbols of observed events.

Motoric reproduction processes are the third component of modeling. The observer must possess the requisite skills for physically executing the behavioral pattern.

Finally, reinforcement and motivational processes determine whether the learned behavior will be activated into overt expression. Positive incentives facilitate translation from mental acquisition to behavioral performance of matching actions. In particular, observed reinforcement contingencies provide the crucial instigational cues.

Bandura notes that the anticipation of reinforcement also influences the attentional and retentional processes; observers are more attentive to actions of effective models, and they are more likely to code utilitarian model responses. However, such incentives are facilitative rather than necessary conditions; indeed, attention may be attracted by physical means: "One does not have to be reinforced, for example, to hear compelling sounds or to look at prominent visual displays." In one study, Bandura, Grusec and Menlove (1966) found the same level of imitative acquisition from television stimuli regardless of advance notification that correct modeling would be rewarded.

Bandura identifies several types of modeling functions. Observational learning is the transmission of information about ways of organizing component responses into new patterns of behavior. Observers acquire novel response patterns by watching the performances of others via physical demonstrations, pictorial representation, or verbal description. While adults rely on verbal modeling as the preferred mode of response guidance, children who have not developed adequate language skills are more dependent on visual depictions. The varied symbolic modeling portrayals on television constitute an extremely important influence, since children "rarely have to be compelled to watch television, whereas verbal characterizations of the same activities would fail to hold their attention for long. One might also expect observers who lack conceptual skills to benefit less from verbal modeling than from behavioral demonstrations."

The second function involves strengthening or weakening of inhibitions that govern the expression of previously learned responses. The observation of reinforcement to



a model helps determine how behavioral restraints are modified; vicariously punished responses tend to inhibit expressions of similar behavior, while normally prohibited responses that are rewarded (or merely ignored or not noticed) reduce inhibitory constraints.

Response facilitation is the third major function. This modeling enhancement process occurs when the model performs a socially sanctioned behavior which serves as an external reminder eliciting performance of existing responses in the same general class. Facilitation is distinguished from observational learning in that new behaviors are not acquired, and from inhibition/disinhibition in that these types of behavior are not normally subject to internal or external restraint.

Bandura briefly mentions a fourth function, stimulus enhancement, where the observer's attention is directed to the objects employed by the model. Consequently, the observer may be more likely to use these objects, although not necessarily in a directly imitative fashion.

In his theory, Bandura posits that human functioning is responsive to the regulatory control of stimulus and reinforcement cognitive processes. Stimulus control is generated by the characteristics of the environmental cues, such as intensity, relevance, explicitness, and model attributes. Stimuli acquire response-directive properties when they are associated with differential response consequences. Such information is conveyed by verbal communication, actions of models, distinctive places, people and things, and pictorial materials. Bandura states, "Of the numerous cues that influence how people will behave at any given moment, none is more ubiquitous or effective than the actions of others," and ". . . behavior is prompted and channelled by the power of example."

Modeling stimuli are not equally influential in evoking the exemplified behavior. The attributes of the social model combine with the model's reinforcement contingencies to determine impact. Models who are perceived to have high competence, expertise, power, celebrity standing or socio-economic status are overtly imitated to a considerably greater degree, compared to models lacking these qualities. Impact of properties such as age, sex, and ethnic status is more likely to vary according to the observer's characteristics, since perceived similarity to the model is an influential factor. The potency of these dimensions is explainable in terms of the predictably reinforcing outcomes associated with imitation of respected or successful models; copying actions of models who have attained status is more likely to be rewarded. Of course, a mass medium such as television often confers status on individuals who appear in messages, since the audience perceives that those dealt with in the media are important. Status also tends to be attributed to models as a result of their response consequences; a model assumes greater valuation and emulative qualities when positively rather than negatively reinforced. In addition, source generalization processes tend to extend a model's prestige to unrelated domains of behavior and to other models similar to the respected one. The role of similarity is partly due to the expectation that people who possess similar characteristics share many common experiences and outcomes. Berkowitz (1962) observes that the degree of similarity between the circumstances portrayed and the observer's own situation facilitates effects in the same manner.

Cognitive control is a key intervening process which helps determine what is observed and performed. Since behavior is primarily regulated by anticipated consequences of prospective overt actions that may not be relevant until appropriate circumstances arise, the cognitive function serves to preserve the imagery and symbolic representations. Thought control is often exercised

before performance; alternative courses of action are tested symbolically, with action based on expected consequences.

Reinforcement control is centrally important, of course. Behavior is extensively governed by its positive and negative consequences, whether directly experienced, observed, or self-created. Actions are shaped by feedback from extrinsic rewards and punishments, more typically those of a secondary social nature (i.e., approval, attention, affection) rather than tangible primary incentives. Intrinsic self-reinforcement also serves as a continuous guide to behavior, as people seek to achieve personal satisfaction by meeting internally evaluated standards.

Observed reinforcement contingencies affect behavior in the same way as directly experienced reinforcers. Bandura argues that vicarious reinforcement is more effective in acquisitional learning, while direct reinforcement is more influential in motivating performance of behavior. A particularly interesting problem is the interactive effect of direct and vicarious reinforcers. Observed consequences provide reference standards for determining whether attained outcomes are judged as positive or negative, as in the phenomena of relative deprivation: "Through social comparison processes, observation of other people's response outcomes can drastically alter the effectiveness of direct reinforcements."

Bandura explains vicarious reinforcement in terms of several mechanisms. One is the informative function, where selective rewarding or punishing of the observed performer's responses conveys knowledge of the probability that certain types of actions are appropriate or successful in various environmental settings. These observed consequences may also provide an incentive motivation function as the observer develops expectations that similar reinforcement can be personally obtained for analogous performances. The specific disinhibitory effect derives from observation of a model engaging in disapproved behavior without experiencing adverse outcomes; inhibitions are released by the perception that such behavior is acceptable in certain situations, and that there is little risk of reprimand or loss of self-respect.

#### Applications

Most of the mass communication applications of vicarious social learning theory have concerned the impact of television violence portrayals on anti-social behavior (see reviews by Bandura, 1973; Liebert, Davidson and Neale, 1973; Atkin, Murray and Nayman, 1971), with more recent research examining the role of modeling processes in learning of pro-social behavior (i.e., Stein and Friedrich, 1972) and cognitive skills (i.e., Lesser, 1974).

The study most relevant to commercial advertising impact is a social learning analysis of public service announcements by Liebert, Sprafkin and Poulos (1975). They obtained successful results by carefully designing "Cooperation" PSA messages to attract attention, produce comprehension, and engender acceptance as a behavioral guide. Based on their experience, the authors recommend these rules for message construction: defining target audience clearly, highlighting central-message while varying peripheral content across a series of spots, using familiar and relevant experiences to optimize immediate recognition and understanding, utilizing action to attract attention and interest, employing verbal labeling cues to emphasize action sequences, and emphasizing strong positive consequences of the behavior.

The remainder of this paper will present recent empiri-

cal evidence pertaining to vicarious modeling, drawn from data collected in a three-year research program by Atkin (1975). In each section, research findings will be linked to relevant components of social learning theory. First, the stimulus characteristics of children's commercials are examined in terms of effectiveness potential, based on social learning ideas. Then, evidence of actual advertising impact is presented in this theoretical context. (The findings will be summarized briefly; full technical reports are available from the author.)

#### Dimensions of Advertising Stimuli

The nature of the portrayal of product consumption and consequences in commercial messages has significant implications for child modeling of advertising stimuli. A number of content dimensions relating to social learning theory were described in an analysis of all 470 commercials aired on the three television networks on two comparable Saturday mornings in 1972 and 1973. In general, these ads portrayed one or more models in the act of consuming a food or toy product and enjoying positive reinforcement as a consequence of this response. The presentation of findings is organized according to the stimulus control and reinforcement control functions, which affect attentional, retentional, and motivational processes.

Stimulus control. The structural elements of advertising presentations contribute to the viewer's attention and retention: ads are placed within and between popular entertainment programs where children have a high opportunity to encounter the messages, and the frequent repetition of specific commercials increases the likelihood of attending and retaining the material. The intermittent repetition is particularly important because younger viewers have a limited capacity for storing an infrequently presented message until the appropriate time for action; since most ads are oriented toward response facilitation effects, repeating the basic message within and across Saturday mornings serves as a key reminder to display previously learned consumption behaviors. The findings showed an average of 20 ads per hour in these two years, and an ad was presented an average of 1.4 times per morning.

The repetition factor is also apparent within commercials. Almost every commercial mentioned the brand name more than once, with almost one-fourth of the ads featuring five or more repetitions. On the average, the name of the brand was verbally mentioned 3.7 times per 30-second message. Retention can also be stimulated by the use of catchy slogans and jingles. More than two-fifths of the ads employed these devices, with musical jingles predominating over verbalized slogans by a five-to-one margin.

Several other attention-getting aspects of advertising stimuli were documented: a humorous tone of presentation was used in almost three-fifths of the ads, and four-fifths of the commercials featured special effects techniques, typically close-ups of the product. Certainly the various model characteristics and reinforcement portrayals described below also induce greater attention.

Almost all ads depicted human characters, and the roles satisfied both the competence/authority and similarity strategies for eliciting that modeling. Nine-tenths portrayed child models, which large segments of the child audience are likely to perceive as similar to themselves. Although few commercials presented adults alone, more than one-fourth had a mixture of children and adults; most adults occupied a parental role. Furthermore, celebrity models were used in one of every twenty ads. Aside from age of characters, a racial ana-

lysis showed that about one-fifth of the commercials presented non-white minority persons, almost always in combination with whites.

Another important facet of the stimulus is the explicitness of the behavioral portrayal. Almost nine-tenths of the ads visually illustrated the product in use at some point, typically showing a child playing with a toy or eating a food product. About three-fourths of all commercials relied solely on live-action portrayals, while the rest were split evenly between animated-only presentations and use of the two techniques in combination.

Reinforcement control. The content analysis also attempted to assess the types of rewards attained by models, although this requires a subjective evaluation. Only a handful of commercials did not portray explicit enjoyment or benefit experienced from consuming the product. Almost three-fifths of the ads were judged to show high satisfaction, with verbal and/or nonverbal display of strong liking for the product; the remainder were rated in the moderate satisfaction category. Higher levels of satisfaction were depicted much more often in food ads than in toys ads. Among the rewards associated with product consumption, the feeling of fun was emphasized in two-thirds of the advertisements, particularly ads for foods. Two other categories were frequently found in toy ads: feelings of power and feelings of being grown-up. Finally, the commercials were rated in terms of peer appeals linked to product consumption. Increased status with peers was demonstrated in very few commercials, but more than half of the ads were classified as containing affiliative appeals since the model(s) and product were shown in a social setting with other children present.

The previously cited data that most adults appearing in ads occupied a parental role with respect to child models has important reinforcement implications. Such authority figures either directly approve of the child model's consumption or implicitly endorse it with their mere presence; this might indicate to the viewer that he will similarly be rewarded (or at least not punished) for consuming the product.

The presentation of such reinforcements in commercials contributes to the cognitive and behavioral impact of the message. Attention and retention are heightened by the enticing display of personal and social satisfaction derived from playing with a particular toy or eating a certain cereal or candy. Most significantly, these elements of the advertisement serve as a motivating incentive to perform the consumption behavior and to implement any necessary pre-consumption actions such as requesting purchase of the product.

#### Attention to Advertising

The initial requisite condition for successful behavioral modeling is securing attention of the target audience. Data from both experimental and survey studies show that commercials attract an attentive reaction from children.

In the experimental research, 500 preschool and grade school students were unobtrusively monitored as they watched a videotaped cartoon in which seven commercials were embedded. Groups of four children were seated in front of a television while a hidden camera recorded their eye contact with the televised presentation. Averaging across the seven 30-second ads, the mean amount of eye contact was 25 seconds. Older children (8-10 years old) paid slightly closer attention than younger children (3-7 years old), by about a 2-second margin.

One of the manipulations in this experiment was the structural formatting of the commercials into a single

cluster of seven ads versus conventionally dispersed sets of three or fewer ads. Attention level was actually slightly higher in the clustered condition, regardless of age. Moreover, there was no significant tendency for relative deterioration of attention from earlier to later ads within the cluster. While these findings underline the attention-drawing power of television ads, it must be noted that the generalizability is somewhat limited by the artificial nature of the laboratory setting.

These experimental data are complemented by survey evidence from a sample of 775 children in the fourth through seventh grades. A questionnaire displayed pictures and/or verbal summary descriptions of 26 specific commercials and PSA's representing a wide variety of products and ideas aimed at both child and adult audiences. Respondents were asked to report level of exposure to each along a four-step scale, with the standard question: "When this commercial comes on TV, how much do you watch it?" Averaging across all messages, 17% selected the "always" category, 24% indicated "usually," 40% said "sometimes," and 19% marked "never." The fourth-fifth graders attended slightly more closely than the sixth-seventh graders. By type of message, attention was reported to be highest for PSA's (53% in top two attention categories), followed by candy ads (50%), hygiene ads (41%), cereal ads (39%), shoe ads (38%), toy ads (33%), and medicine ads (25%). Of course, the particular examples of ads selected for each type of product may not be representative, but the overall findings are likely to be typical.

#### Observational Learning Effects

Observational learning is defined as the acquisition (and performance) of novel forms of behavior. In the case of television advertising, the criterion of "new" response can be met with any of several conditions including: a type of consumption with which the child is not familiar because of limited opportunity for direct observation or experience (i.e., using hygiene products such as deodorants and acne cream, ingesting proprietary medicines such as sleeping pills and laxatives, wearing undergarments, or drinking alcohol), or a distinctive new version of familiar type of consumption (i.e., playing with unique style of toy, or eating an innovative brand of dessert). Several sets of data are relevant to observational learning:

Hygiene socialization. In the survey described in the previous section, respondents described their attention to commercials for deodorants, mouthwashes, and acne creams, and their orientations toward these products. Among the criterion variables studied, three are most likely to be affected by modeled stimuli portraying numerous actors worrying about hygiene and using advertised products to combat the problem: perceived amount of usage of the products by other people, personal concern about the problems, and usage of the products.

The predictor variable was a multiplicative exposure index composed of degree of attention to five hygiene ads and amount of television viewing during prime-time evening hours when these ads are aired. Three items measured the children's perceptions of the proportion of adults (or teenagers, in the case of skin cream) who use each of the hygiene products. An index based on these items correlated +.30 with the exposure index; when grade, sex, social status, scholastic performance and interpersonal discussion about hygiene were controlled, the partial correlation remained sizable at +.26. Thus, heavy hygiene commercial viewers are much more likely to perceive that people frequently use deodorants, mouthwash, and skin cream.

The motivational element was gauged by the extent to

which the child worried about acne problems and offending others with body odor. This index correlated +.20 with exposure, and the fifth-order partial correlation dropped to +.14. There was correlation of +.29 between exposure and the frequency of actually using the mouthwash and skin cream products; the partial correlation was +.23. These findings indicate a moderate association between viewing hygiene ads and using these products, and a somewhat weaker relationship for concern about hygiene problems.

The effect of advertising was experimentally tested in the case of acne cream. In a sample of 200 fourth and fifth graders, half viewed a video tape program containing a currently aired acne cream commercial while the other half saw an irrelevant commercial in that slot. The message showed pre-adolescents applying the preparation to their faces. On a post-viewing questionnaire, the exposed group expressed significantly greater concern about blemishes, were significantly more likely to believe that skin cream rather than regular soap was the appropriate solution to acne problems, and were significantly more likely to say that they would buy the product for personal use, compared to the control group. A contingent analysis showed the importance of the novelty factor in message impact: among the 83 subjects who had previously seen the product advertised on TV, there was no difference between the experimental and control groups. However, among those who had never viewed such messages at home, the difference was very large, yielding highly significant interactions between the treatment variable and the previous exposure variable.

Point-it-out. One unique behavior that has been emphasized in a well-known anti-pollution PSA is to point-it-out when polluting or littering is encountered. A series of models distinctively use their fingers to identify various offenders and perhaps to stimulate them to stop polluting. Since children would be unlikely to perform this response without observing the PSA stimulus, it is a prime example of observational learning. In survey questionnaires distributed to 775 older children, one question asked how often they reminded litterers to stop their littering; amount of exposure to this and another anti-littering message was also measured. More than two-thirds of the respondents indicated that they had ever attempted to stop offenders from littering; 25% of the heavier viewers vs. 14% of the lighter viewers reported that they did this frequently. The partial correlation between exposure and reminding others to stop littering was +.19, with demographic factors controlled. It appears that the novel act of harassing litterers is effectively taught by televised PSA's.

Model characteristics. Among the stimulus control variables studied in the social learning literature, much attention has focused on model attributes such as sex, age, and race. The racial identity of the advertising model was examined in one of the experiments of this project.

Previous research yields conflicting results. Neely, Heckel and Leichman (1973) found that young black children imitated a televised white model more than a black model in toy selection. However, Barry and Hansen (1973) reported that young black children were much more likely than whites to express preference for a cereal brand promoted in a TV commercial featuring a black model than an alternative brand advertised without a black model.

In the present experiment, the race of models was manipulated in a specially produced commercial for an unfamiliar toy product. The ad presented a pair of either white actors or black actors playing with and enjoying the toy; all other message elements were held constant.

Half of the sample of 500 preschool and elementary school children viewed each condition. There were no major differences in observed reactions while viewing, in terms of attention, irritation, enjoyment, and verbalizations. In a post-viewing play situation, subjects were offered an opportunity to play with either the advertised toy or an alternative toy. For white subjects, there were no differences in selection between the white-model and black-model treatments; black subjects more often chose the advertised toy in the white-model condition by a clear 14% margin. These findings indicate that black children are somewhat more likely to imitate novel responses of racial dissimilar advertising models.

#### Inhibitory/Disinhibitory Effects

In addition to teaching new responses, commercials may modify inhibitory restraints governing performance of existing behavioral patterns. Although previous television applications of this component of social learning theory have generally dealt with clearly proscribed forms of behavior such as aggression, the advertising derivations pertain to disinhibition of more mildly prohibited behavioral expression involving minor violations of personal standards or social norms, including: counter-normative types of product consumption (i.e., unconventional definition of appropriate product users or situations for product usage, such as girls playing with electric trains or late-evening snacking with breakfast cereal), and atypical amounts of product consumption (i.e., candy eating beyond normal limits).

Since commercial advertising seldom attempts to restrain behaviors by presenting negative reinforcement of product consumption, inhibitory applications are rare. However, some public service campaigns seek to inhibit the display of generally non-sanctioned forms of behavior such as littering, smoking, drug-taking and car-riding without buckled seat belts. While the goal may not be total prohibition, PSA's often endeavor to restrict the frequency or situational conditions for performing such behaviors.

The research project included evidence dealing with both disinhibition and inhibition of acquisition and performance, following from the examples listed above.

Recreational sex role socialization. Modeling influences may serve to teach non-traditional forms of play through advertising. To test the impact of a counter-stereotypical use of a toy racing car set by female models, two versions of a standard toy commercial were filmed: one portrayed two young boys playing with the racing cars, while the other presented the same actions acted by two young girls. Half of a sample of 400 second through fifth graders viewed each version, which was inserted into a cartoon program on video tape. On a post-viewing questionnaire, the manipulation check disclosed an interesting finding: just two-fifths of the subjects exposed to the female-model condition recognized that they were girls. The dependent variables were the perceived appropriateness of girls playing with racing cars and personal desire to play with the toy. There were no significant main effects between the two experimental treatments. However, examination of the subgroup that correctly perceived the sex of the female models showed substantial differences in the expected direction: this subgroup was almost twice as likely to feel that girls should play with racing cars, and they were slightly more likely to indicate that they would like to play with the cars (although they did not express a greater intention to actually ask parents to buy the expensive toy). However, these latter findings may be partially due to self-selection of more attitudinally favorable subjects into the accurately perceiving subgroup. The pattern of results did not differ accord-

ing to sex of the subject.

Candy consumption disinhibition. Since almost all children eat candy products at least occasionally, moderate consumption is hardly a novel or proscribed pattern of behavior. Nevertheless, Saturday morning candy commercials repeatedly portray models happily consuming a variety of these candies, often with the additional reinforcement of tacit adult approval. Extensive exposure to these modeling stimuli may suggest to the child that excessive candy eating is acceptable behavior, even though interpersonal messages and intrapersonal judgment typically serves to restrain candy intake. Thus, commercials may cause a reduced level of personal guilt or fear of social disapproval for excessive consumption of candy; this effect should be reflected in greater amounts of candy bars eaten by the child. Since inhibitions probably do not pertain to any particular brand of candy, such an effect should be generalized to consumption of all brands regardless of the frequency that each is advertised on television.

In a questionnaire survey of 506 fourth through seventh graders, exposure to candy advertising was indexed by a generalized candy attention item and two measures dealing with specific ads, multiplied by the amount of viewing time on Saturday mornings. Children were also asked to report how often they ate seven brands of candy bars, including three that were frequently advertised and four that were lightly advertised on Saturday mornings. Exposure correlated +.29 with an index of eating heavily promoted brands, and the partial controlling for demographics was +.25. Consumption of lightly advertised brands was associated with exposure to the same degree. A separate item measuring number of candy bars consumed per week yielded a weaker partial correlation of +.10 with the exposure index. Thus, there is fairly strong evidence that exposure and consumption are functionally related (although the direction of causality is not unambiguous) and that advertising effects tend to be generalized to nonadvertised brand consumption, lending support to the disinhibition explanation.

PSA's and inhibitions. Social learning theorists argue that young people learn to avoid dangerous mistakes by observing the negative consequences incurred by models who don't follow recommended practices, such as buckling seat belts or restricting smoking, rather than learning by direct experience with punishing outcomes. Theoretically, an effective modeling message should graphically display a model performing the illicit behavior and then suffering injury, illness, or death. Public service announcements typically tone down the explicitness of the harmful consequences, but generally follow this basic approach. Two cases are seat belt and anti-smoking campaigns, which display such negative reinforcement as hospitalization, loss of affection, or physical disability.

The questionnaires administered to 775 older children contained items dealing with seat belt buckling behavior and intention not to smoke, along with level of exposure to corresponding public service messages.

Controlling for contaminating demographic variables, there was a negligible +.06 partial correlation between seat belt PSA exposure and frequency of using seat belts; the partial correlation between exposure to anti-smoking PSA exposure and intention not to adopt cigarette smoking was -.11, however. It appears that these PSA's have no positive impact in increasing inhibitions against smoking or riding without seat belts, probably due to the extensive amount of interpersonal influence on these behaviors from parents, teachers, and peers.

One particularly pervasive public service campaign has sought to restrain littering behavior. The negative

consequences of this type of action are qualitatively different from the previous examples, since the harmful outcomes are aesthetic or social rather than physical. PSA's typically portray littering responses followed by portrayals of a polluted environment, social disapproval to the model, or a saddened Indian. In the survey, exposure to this type of message was associated +.05 with frequency of proper disposal of litter, with demographic factors controlled.

An experimental test examined the impact of the crying Indian spot on actual littering behavior under controlled conditions. Half of the sample of 500 subjects in the previously discussed experiment viewed this anti-littering message during an entertainment program, while the other half were exposed to an unrelated message. After the viewing session, all subjects were offered a piece of wrapped candy and their disposal of the wrapper was unobtrusively observed by the experimental assistants. The children who viewed the littering announcement were significantly less likely to litter the experimental room than the non-exposed Ss. Among viewers, 25% put the wrapper in the waste basket and 2% threw it on the floor; 19% of the non-exposed Ss disposed of the wrapper in the basket and 11% littered the floor. The others kept it on themselves or left it on the table. The younger children were clearly most influenced. While the experiment demonstrates at least a temporary inhibitory effect, the survey findings show that the influence is not strong in the naturalistic setting.

#### Response Facilitation Effects

The response facilitation function is highly relevant to modeling of advertising stimuli. Most product consumption in commercials is acceptable, everyday behavior encumbered by minimal restrictions. The same holds for certain socially desirable practices promoted in public service announcements. For instance, children have established habitual patterns of behavior regarding cereal eating, toy playing, and in most instances, litter disposal and seat belt buckling. They have typically experienced direct positive reinforcement for such actions, but performance may not be salient at any given moment. In such cases, advertising might serve as a discriminative cue instigating previously learned behaviors; the ads temporarily remind the child to perform the particular activity. These eliciting cues typically contain a positive rather than negative motivational component. In particular, commercial messages seldom depict the unfavorable consequences of not using the product (except in preliminary scenes before the product is consumed, as in medicine and hygiene ads); the primary modeling sequence is the demonstrated use of the product followed by rewarding outcomes.

An important aspect of this process concerns stimulus and response generalization from a given message. Does advertising motivate consumption of any brand within a product class, or is the effect restricted to the specific brand presented in the commercial? For instance, an ad for one exotically flavored pre-sweetened cereal may trigger eating of similar sugary cereal brands, or the generalization gradient may further carry over to non-sweetened cereals. This generalization process might be expected with children for two basic reasons: the brand-unique cues in a given ad might be perceived as peripheral to the more central modeled sequence of preparing and eating a bowl of cereal (or playing a board game or using a hair shampoo), and the overall message environment of numerous competing cereal brands (or game brands or shampoo brands) may produce a cumulative impression of "eating cereal" (or "playing games" or "using shampoo") rather than learning substantive and symbolic distinctions between brands X, Y, and Z. From a visual modeling perspective, the basic theme is consumption behavior; there are often no unique features of

specific brands that models can act out (except listening to snap/crackle/pop). Of course, some models may be more influential exemplars, and there may be non-modeling aspects such as nutritional value, taste qualities, and attractive brand symbols. These other factors should create a tendency for children to more frequently consume heavily advertised brands--but a substantial generalization to consumption of less advertised brands should also be expected.

This notion was tested in a survey investigation of cereal advertising and consumption, where 506 children in the fourth through seventh grades reported on their exposure to cereal ads (indexed by the product of five specific and general attention items times amount of Saturday television viewing) and their patterns of cereal eating. A consumption index for eight heavily advertised cereal brands correlated +.41 with the exposure index; when grade, sex, status and school performance were controlled, the partial correlation remained strong at +.37. As anticipated, the association between exposure and consumption of five lightly advertised brands was positive but less strong, with a raw correlation of +.27 and a fourth-order partial of +.24.

Since response facilitation might be counteracted by parental influences restricting the amount of cereal eating, the questionnaire also measured whether parents imposed general snacking rule limitations or allowed children to eat whatever foods they pleased. The conditional correlations showed a substantial interaction: in homes with no rules, the correlation between exposure and overall cereal consumption was +.51, while the correlation was only +.27 for children reporting parental restrictions.

While it is possible for children to model cereal eating with the cereals available in the home, imitation is to some extent contingent on the intervening variable of persuading parents to purchase cereals in the supermarket. The motivation to imitate may indirectly produce request behavior as a means to physically acquiring the product (advertising codes discourage modeling stimuli of the product acquisition behavioral sequence of asking parents to buy, which would be a much more effective response facilitation strategy). The survey showed that exposure to cereal advertising was correlated +.32 with frequency of requesting cereal purchases, with a partial correlation of +.27. Of course, it should be recognized each of the relationships reported in this section may be due to reverse causation, as children who often request or eat cereals are motivated to watch TV commercials for cereals. Since a major portion of the exposure variable involves mere amount of time viewing Saturday morning programming, it seems unlikely that a large part of the association could be explained by selective seeking rather than message effects.

Analysis of path coefficients among exposure, requests, consumption and demographic variables indicated that the impact of cereal advertising on advertised brand eating occurs both directly and indirectly via requests. The linkage from exposure to requests was +.27, and from requests to consumption was +.26; the direct path coefficient from exposure to consumption was +.30.

This set of findings suggests that cereal advertising has a direct response facilitation effect on consumption of advertised cereals, and that this impact is somewhat generalized to less frequently advertised cereals as well. Apparently the repeated presentation of models eating cereal serves to remind the child to perform habitual cereal eating behaviors. Furthermore, there are indications of an indirect stimulation of purchase requests as a means to obtaining the product for consumption.

In the case of public service announcements, modeling effects have been primarily interpreted as observational learning and inhibitory processes. For frequently expressed sanctioned behaviors such as seat belt buckling and proper disposal of trash, an alternative approach emphasizing positive reinforcement can be utilized. Many learning theorists believe that reward is more effective than punishment in shaping behavior. The positive facilitation strategy would be to portray the recommended practices paired with rewarding consequences, such as safe arrival, good health, social approval, self-satisfaction, or a clean environment. Since few PSA's use this approach, the data from this investigation do not provide a good test for these ideas; however, the research of Leibert, Sprafkin and Poulos (1975) suggests that it can be effective.

#### Stimulus Object Enhancement Effects

Bandura suggests that stimulus enhancement processes direct the observer's attention to objects utilized by the model and produce greater use of the objects beyond demonstrated circumstances. This notion can be applied to evidence on medicine advertising effects gathered in a survey of 256 fifth, sixth and seventh grade students. The analyses showed that exposure to advertisements of headache remedies was slightly associated with the child's self-prescription of the number of aspirin that they should take to relieve a cold; the partial correlation was +.06, with 16% more of the heavy medicine ad viewers than light ad viewers suggesting a dosage of two or more aspirin. On the other hand, there was no difference according to exposure on an item asking the children if it's acceptable to take aspirin if they're not really sick. This evidence provides very limited indications that an enhancement of advertised objects can occur for a significant behavior such as medicine taking.

#### References

- C. Atkin, "The Effects of Television Advertising on Children," final report submitted to Office of Child Development (1975).
- #1 First Year Experimental Evidence (June, 1975).
  - #2 Second Year Experimental Evidence (June, 1975).
  - #3 Exploring the Relationship between Television Viewing and Language Development (June, 1975).
  - #4 Attitudes of Industry Executives, Government Officials and Consumer Critics Toward Children's Advertising (June, 1975).
  - #5 Content Analysis of Children's Television Commercials (June, 1975).
  - #6 Survey of Pre-Adolescent's Responses to Television Commercials (July, 1975).
  - #7 Parent-Child Communication in Supermarket Breakfast Cereal Selection (October, 1975).
- C. Atkin, J. Murray, and O. Nayman, Television and Social Behavior: An Annotated Bibliography of Research Focusing on Television's Impact upon Children (Washington, D. C.: U. S. Department of Health, Education, and Welfare, 1971).
- A. Bandura, "Social Learning Through Imitation," in M. Jones (ed.), Nebraska Symposium on Motivation (Lincoln: University of Nebraska Press, 1962), 211-274.
- A. Bandura, "Vicarious Processes: A Case of No-Trial Learning," in L. Berkowitz (ed.), Advances in Experimental Social Psychology, Vol. 2 (New York: Academic Press, 1965), 1-55.
- A. Bandura, "Social-Learning Theory of Identificatory Processes," in D. Goslin (ed.), Handbook of Socialization Theory and Research (Chicago: Rand McNally, 1969), 213-262.
- A. Bandura (ed.), Psychological Modeling: Conflicting Theories (Chicago: Aldine-Atherton, 1971a).
- A. Bandura, Social Learning Theory (New York: General Learning Press, 1971b).
- A. Bandura, Aggression: A Social Learning Analysis (Englewood Cliffs, N. J.: Prentice-Hall, 1973).
- A. Bandura, J. Grusec, and F. Menlove, "Observational Learning as a Function of Symbolization and Incentive Set," Child Development, 37 (1966), 499-506.
- A. Bandura, D. Ross, and S. Ross, "Imitation of Film-Mediated Aggressive Models," Journal of Abnormal and Social Psychology, 66 (1963), 3-11.
- A. Bandura, D. Ross, and S. Ross, "Vicarious Reinforcement and Imitative Learning," Journal of Abnormal and Social Psychology, 67 (1963), 601-607.
- A. Bandura and R. Walters, Social Learning and Personality Development (New York: Holt, Rinehart and Winston, 1963).
- T. Barry and R. Hansen, "How Race Affects Children's TV Commercials," Journal of Advertising Research, 13 (1973), 63-67.
- L. Berkowitz, Aggression: A Social Psychological Analysis (New York: McGraw-Hill, 1962).
- J. Gewirtz, "Mechanisms of Social Learning," in D. Goslin (ed.), Handbook of Socialization Theory and Research (Chicago: Rand-McNally, 1969), 57-212.
- L. Kohlberg, "Stage and Sequence: The Cognitive-Developmental Approach to Socialization," in D. Goslin (ed.), Handbook of Socialization Theory and Research (Chicago: Rand-McNally, 1969), 347-480.
- G. Lesser, Children and Television: Lessons from Sesame Street (New York: Random House, 1974).
- R. Liebert, J. Neale, and E. Davidson, The Early Window: Effects of Television on Children and Youth (Elmsford, N. Y.: Pergamon Press, 1973).
- R. Liebert, J. Sprafkin, and R. Poulos, "Selling Cooperation to Children," in W. Hale (ed.), Proceedings (New York: Advertising Research Foundation, 1975).
- N. Miller and J. Dollard, Social Learning and Imitation (New Haven: Yale University Press, 1941).
- J. Neely, R. Heckel, and H. Leichtman, "The Effect of Race of Model and Response Consequences to the Model on Imitation in Children," Social Psychology, 89 (1973), 225-231.
- B. Skinner, Science and Human Behavior (New York: Macmillan, 1953).
- A. Stein and L. Friedrich, "Television Content and Young Children's Behavior," in J. Murray, E. Rubinstein, and G. Comstock (eds.), Television and Social Behavior, Vol. 2 (Washington, D. C.: Government Printing Office, 1972).

## EVALUATING CHILDREN'S RESPONSES TO TELEVISION ADVERTISING

Barbara R. Fowles, Children's Television Workshop

## Abstract

The claim that television viewing is a passive activity has several possible implications. Passivity can mean physical inactivity, taking in preformulated information, abdication of judgement or relinquishing of attention control to a dynamic stimulus display. Only in the latter instance is the term "passive" relevant to the interaction between child and television. The implications of this for assessing the impact of television are discussed.

America has been hawking goods to its children for a long time, and no one has worried about it very much. Recently this state of affairs has, for many reasons, begun to change. Television has been a primary object of concern, in part because the enormous investment of money and air time in advertising directed at children has come to public attention, and partly because our suspicions that children, particularly pre-schoolers, watch enormous amounts of television have been confirmed by research. (C.F. Murray, 1973) Finally, educational programs, Sesame Street in particular, have caused parents and teachers to notice, sometimes with astonishment, that very young children absorb a great deal from apparently casual television viewing. Thus what was once considered a minor factor in the ecology of childhood has become a force to be reckoned with.

The fact that we have now come so far as to consider theoretical approaches to the study of this phenomenon is noteworthy. It means that our empirical approaches to the study of the impact of television on children will be increasingly informed by a broader theoretical perspective; but it also means that our theories of child development may at last begin to reflect what the medium is teaching us about the nature of children's learning.

The following discussion, then, deals with the passivity, a concept that plays a crucial role in the interpretation of responses of young children to television, and which has been wielded to obscure as well as to illuminate those responses. The role of this notion in guiding our observations is explored and evaluated, and the meaning of the concept reconsidered.

When the first television generation was growing up in the 1950's, there was fuss about children becoming physical and intellectual vegetables from watching too much of the stuff. Television, it was argued, demanded a passive response mode which children would then adopt as a general pattern of response to the environment. An important point here is that passivity was not considered an optional response, as it is for the person who stands on the sidelines during a pick-up basketball game, but as required by the dynamic unidirectional character of the medium. Little evidence was forthcoming that these children had noticeably more vegetable-like attributes than their predecessors and the idea of a nation of pale, obese, near-sighted maladjusted children faded largely away. Passivity has nevertheless continued to be ascribed negatively to the young television viewer, though its attribution now has the weight of developmental theory behind it. I will briefly describe several such notions:

According to Piaget, the most widely familiar and accepted of the developmental theorists, children up to about the age of seven, have their thought processes firmly grounded in practical action (Piaget, 1963). In the sensorimotor period which lasts until about two years of age, the child's primary mode of mental growth and development is said to be through direct motor involvement with the physical environment. The child builds a coherent picture of reality by acting on the environment and getting feedback from it.

The two-to-seven-year-old child increasingly carries out these "experiments" in her head instead of through action, but the mental manipulations are still intuitive, and involve visualizations of concrete events. Only at the next stage is this intuitive mode of reasoning replaced by logical structure which is still very closely tied to practical actions. Auditory and visual learning from television is, of course, acknowledged within this theory, but information gotten through the eyes and ears is supposedly incorporated into preexisting cognitive structures (assimilation) rather than initiating change in those structures (accommodation.) Given the Piagetian view of the process of mental growth, television plays a minimal role because it does not allow for direct interaction. It does not provide for the sort of experience with physical and social reality that is the raw material for real intellectual progress. Thus within the context of this theory, the passivity of the viewing child is understood in the following senses: Learning from television is passive in the rather literal sense that the child is not directly engaged with the environment, but merely gets pre-formulated information through the auditory and visual channels. The child is physically and socially unresponsive in her role as television viewer.

Learning from television is passive in the sense that information gotten through perceptual experience, as opposed to practical experience, will not initiate cognitive change in young children, but rather will be comprehended in terms of the existing conceptual structures.

Developmental psychology offers yet another way of looking at passivity. Ability to systematically allocate ones attention, (to look and listen selectively and efficiently) develops gradually. Young children are captivated by external stimuli and easily distracted from one focus of attention by another competing one. As the child grows, she becomes better at imposing efficient, internally guided search strategies on the stimulus field in order to get needed information from the environment. This is not a matter of acquiring control over perceptual mechanisms, but one of cognitive growth. The very young child knows how to look in the mechanical sense, but has no basis for determining what to look for (Vurpillot, 1968), and how to plan and guide her search (Miller, Galanter, Pribram 1960).

For a young child, a dynamic display like television is a particularly powerful controlling stimulus. The eye is coaxed around the screen by light and movement, and presumably the mind goes along for the ride. The view-

er is passive in that while she may be (though is not necessarily) processing information rather rapidly, she is not directing a search for information, but is apparently being "fed" a sequence of percepts dictated by the television screen and its auditory accompaniment.

If the sequence of percepts is largely dictated, then it follows that evaluation of these percepts (for plausibility, truth value, etc.) is largely suspended, since judgemental activity is not a passive mode of response.

So we have collected two more interpretations of the concept of passivity: First, the child is neither systematic nor selective in processing the information television makes available during the course of viewing rather successive foci of attention are largely dictated by the stimulus structure. The TV, not the young viewer, is in charge. Second, this lack of selection would seem to imply passive acceptance of the message.

When it comes to considering the effects of television advertising on children, working through these various formulations of the concept of passivity comes to be particularly relevant because two points of view, with important differences in their implications, are possible. The first position is that since young children learn mostly through direct experience with the environment, the impact of television, negative or positive, is negligible. To extend this to advertising is to minimize the effects of television advertising on young viewers by arguing that the young child is simply not the sort of animal to be much influenced via the communication modes available to that medium. Hence, exposure to television commercials is essentially harmless, and at most, wasteful of the child's time as well as the advertiser's money.

The second and equally logical position has almost opposite implications: The idea here is that since passivity implies lack of selectivity and lack of critical stance toward incoming stimuli, the child is a willing victim to all advertising messages--a sucker. In this view advertising is powerful and insidious.

Thus the idea of the passive relationship between the child and the TV set can be bandied about to support quite different views of the medium's effects, and quite different frameworks for investigating and interpreting identical observable behaviors. Yet the two views are not actually contradictory, there is merely a difference in the impact ascribed to information gotten indirectly from TV rather than directly through the environment. It is therefore worth sorting this out in order to determine in what sense, if any, passivity is relevant to an understanding of the impact of advertising on children. For if we can tell what a "passive" role for the viewing child means, we are better equipped to look in the right places for harmful effects and to predict what they are likely to be.

To sort this out, we need to go back and examine the first set of statements about passivity. We should do this in light of the fact that the theoretical foundations of developmental psychology were pretty well set before television was invented and, (for reasons ranging from conservatism through snobism,) the field has so far taken remarkably little serious notice of children's responses to television. This is not to say that even large doses of television have essentially altered the course of cognitive growth. However, the ubiquity of television has indeed changed the nature of the environment in which those children who are heavy viewers grow up, and it is of course these children who are of primary interest to us. Furthermore, it is probably true that nothing with an impact even approaching

that of television has been introduced into the ecology of childhood since formal education, which had an acknowledged effect on the course of development, became widespread. It seems at least possible that children have always been capable of functioning in ways that were simply not demanded prior to television (and still are not demanded in the range of settings and tasks psychologists typically use to study children.) So developmental psychology's predictions for television ain't necessarily so, and may underestimate both the ability of young children to learn in certain ways, and the power of television to teach them. We need to examine the statements about passivity that have been developed here critically, in light of what we know about television and children's responses to it. We can then go on to identify likely parameters of the impact of television advertising on children which would be appropriate areas for future research.

First and most straightforward is the contention that the young child does not interact with television on the physical plane. In fact young children frequently become actively involved with television programs--dancing, pointing, imitating gestures, singing, and conversing with the television set. Sesame Street has consistently and intentionally provided for this propensity and turned it to desirable ends. This is a matter of designing instructional sequences such that verbal and motor behaviors appropriate to the child's level of development are explicitly modeled. For example, children are encouraged, by an on-screen model, to hold up fingers as they count. And though this sort of response doesn't happen at all possible occasions, and there are individual differences in the tendency to respond, episodes of overt participation are consistent enough to make it quite clear that a physically, even verbally interactive mode of response to television is open to young children. In addition, of course, there is now a considerable body of research which demonstrates that children will imitate behaviors they have seen on TV later on, away from the television. This deferred imitation means that even when no overt response is evident, attention and memory are engaged sufficiently to store the critical attributes of the performance.

The second formulation of passivity restates the first, but on a cognitive rather than an action plane. The claim is that the form of experience television offers makes a trivial contribution to cognitive growth because the lack of motor involvement precludes significant cognitive restructuring. We can, again, question this assumption on two grounds: First, while concept formation in young children probably does require some direct experience, television has obvious ability in contrast to the "real world", to present an organized range of additional examples, to highlight critical features, and to concentrate this experience in time and space. It can therefore accelerate concept acquisition once the experiential basis for the concept has been established. (A verbal formulation of the concept, in addition to examples repeatedly demonstrated that young children can learn concepts like "circle" and "up" from television. These are, of course, concepts which young children encounter in their physical and verbal environment quite frequently, and the interaction between Sesame Street and the environment has no doubt facilitated concept acquisition. Even if television plays the most minimal role of directing the child to aspects of his environment relevant to a particular concept, it is playing a significant role in cognitive growth; and it is almost certainly doing far more than that.

There are even ways in which television material can be organized to give rise to the kind of cognitive restructuring



turing identified with real intellectual growth. According to Piagetian theory (Charlesworth, 1969), a stimulus structure that embodies a perceived conflict (an incongruity, a surprise, a dilemma) will lead an individual to engage in cognitive activity in order to resolve the conflict to her own satisfaction. Of course the stimulus material must be at the level where the viewer can grasp its structure, yet not be bored by it, otherwise, violations of that structure will go unnoticed; hence unresolved. This sort of cognitive activity, which demands reinterpretation of the stimuli, can lead to accommodative activity; or a reorganization of the child's cognitive structures, in order to account for the properties of the stimulus which were initially inexplicable. The significance of this activity depends, then on the importance of the conflict embodied in the stimulus for intellectual growth, and the appropriateness of the stimulus design. We have seen in our research at Children's Television Workshop (Fowles, 1972), stimuli of this kind, particularly events with magical properties, will lead young viewers to generate imaginative hypotheses to explain these properties. Since such stimuli can be designed such that the hypotheses the viewer is likely to produce, covertly and overtly, are (1) in conflict with her currently held notions of how reality works and (2) advanced and adequate as explanations as well, a favorable restructuring of the conceptual system is a likely outcome of exposure to some amount of such stimulus material.

Once we look beyond the surface attributes of the medium and consider stimulus structure, it is rather easy to understand that the so-called "passive" medium can be rendered active by engaging the viewer to resolve or complete the structure.

For either of these mechanisms to work, however the content of the stimuli must of course be appropriate to the child's level of development (Fowles and Voyat, 1975) so that the child has relevant expectations. Deliberateness in the stimulus design is thus required to make the most of this property, yet the contention that television viewing is limited as a learning experience is nevertheless clearly untrue.

The final sense in which the notion of passivity is used is to identify the situation wherein the allocation of attention within the stimulus (primarily the scan path) is controlled by aspects of the stimulus itself, rather than by a search strategy imposed by the viewer. Eye-movement recording (O'Bryan, 1972) has demonstrated that this kind of passive response does occur, especially in young children and children with learning problems. That is, records of children's scan patterns while viewing television shows that they are predictably drawn to certain aspects of the visual display, and as the child matures, an organized scanning strategy, reflecting the desire to get specific information from the screen, emerges.

In this sense television viewing can legitimately be called passive. For instructional purposes this is not necessarily bad, for if the scanning sequence can be controlled, then at the same time a useful strategy for scanning other similar stimuli, outside of the television context, is being modeled for the viewer. It has been proposed (Salomon, 1972) that such scanning strategies may be internalized as thinking strategies, and there is some empirical evidence that this occurs. For example, subjects who were shown scanning strategies modeled by the camera as it examined complex paintings, later displayed similar strategies in their own eye movements, in freely examining still slides of stylistically similar paintings.

Passivity in this sense is thus a meaningful concept

when applied to television. It also has a corollary as noted earlier, which is that in relinquishing control of one's attention strategy evaluative judgement of the stimulus and its message is also relinquished. The latter claim is difficult to document, but is clearly a deserving focus of concern. However, we cannot go so far as to assume, even if we can empirically determine that a particular piece of television material will reliably control subjects' visual attention patterns in a particular way, that a predictable message is thereby being conveyed. Again, the child's level of cognitive functioning determines how information will be interpreted, and for the young child this interpretation may vary widely from the intended message. Thus the concept of the viewer as an information sponge is also something of an erroneous one.

Several points emerge, I think, from looking closely at what passivity means in terms of the child as television viewer in general, and viewer of television advertising in particular. We have seen that the interaction between child and television is complex, and there is no sense in which we can simply say that children are passive in their relation to television, and that they are therefore either wholly immune or wholly vulnerable to what they see. So while we clearly cannot ignore the impact of television advertising, neither can we predict effects from examining the commercials alone. Rather, the precise nature of the child's response to television is dependent on both developmental factors in the child and structural factors in the commercial or program. When these factors are well meshed, television can and often does provide a full range of learning experience for good or for ill. Advertising, perhaps often only by chance, surely participates in this instruction process and needs to be examined from this point of view.

#### References

- William Charlesworth, "The Role of Surprise in Cognitive Development," In D. Elkind and J. Flavell (Eds) Studies in Cognitive Development: Essays in Honor of Jean Piaget, (N.Y.: Oxford University Press, 1969).
- Barbara Fowles, "A Pilot Study of Verbal Report in Formative Research in Television," Unpublished Doctoral Dissertation (N.Y.: Yeshiva University, 1972).
- Barbara Fowles and Gilbert Voyat "Piaget Meets Big Bird: Is TV a Passive Teacher?," Urban Review, (Vol. 7, 1974).
- George Miller, Eugene Galanter, Karl Pribram, Plans and The Structure of Behavior, (N.Y.: Holt, 1960).
- James, Murray, "Television and Violence: Implications of The Surgeon General's Research Program," American Psychologist, (June, 1973).
- Kenneth O'Bryan, "Report on Children's Viewing Strategies," Unpublished Report, (Ontario Institute for Studies in Education, Toronto, 1972).
- Jean Piaget, Six Psychological Studies, (New York: Random House, 1963).
- Gavriel Salomon, "Can We Effect Cognitive Skills Through Visual Media? An Hypothesis and Initial Findings," AV Communications Review, (1972, 20, (4) ) 409-422.
- Emile Vurpillot, "The Development of Scanning Strategies and Their Relationship to Visual Differentiation," Journal of Experimental Child Psychology, (1968) 632-650.

VISUAL AND VERBAL MEMORY IN CHILDREN'S PRODUCT  
INFORMATION UTILIZATION<sup>1</sup>

John R. Rossiter, University of Pennsylvania

Abstract

Consumer research has largely ignored product-relevant information stored in visual memory. Visually stored images may be quite sufficient to engender product choice, quite apart from attitudes, beliefs, and so on, retrieved from verbal or symbolic memory. Such may be the case when children or adults encounter familiar products on a supermarket shelf: visual recognition latency may be all that is involved in the choice process. The present experiment (1) confirmed the existence of a rich data base in children's visual memory for cereal brands, (2) demonstrated that visual information differs from verbal information and that children's choices may differ depending on which of the two types of information is situationally retrieved, and (3) pointed out measurement biases favoring verbal memory in contemporary consumer research and urged more research on visual memory, both with child as well as adult consumers.

Introduction

Most consumer choice models based on the information processing approach presume that choices derive from a synthesis or integration of beliefs (pertaining to information about the choice object) which are largely verbalizable. That is, the models assume that beliefs are represented linguistically in memory. Thus we have our Ss exhibit their belief probabilities, belief evaluations, and so forth, via questionnaires and rating scales. These procedures rely almost exclusively on verbal task instructions.

The author and his colleagues (Calder, Robertson and Rossiter, 1975) have suggested that this reliance on verbally encoded information may overlook a great deal of nonverbal information which consumers utilize in decision making. One important type of nonverbal information is visual information. Its importance seems to be increasing in our culture with the dominance of television as an information medium, particularly as a medium for the advertising of products to children.

This study examines children's utilization of visually stored information about cereals. The general hypothesis is that children have acquired and stored in memory a rich visually encoded data base on cereal attributes. This type of data representation may differ totally from any verbal representation of cereal attributes. Indeed, it may not be amenable to verbal retrieval because a translation from one code to another would be required in the output process (see, for example, Paivio, 1971; Posner, 1973). This transformation may be beyond the capacity of young children.

<sup>1</sup>The study was conducted at The Wharton School's Center for Research on Media and Children. The author wishes to acknowledge the support of Professor Charles S. Goodman, Chairman of Wharton's Department of Marketing, and to thank John Trent for assistance in the research administration.

The importance of studying children's visual memory has not yet become apparent in children's consumer research. Its importance lies in the fact that retrieval of visual representations of (for example) cereal attributes may, alone, be sufficient to engender product preference. The child need only retrieve a visual match for the cereal in a "familiarity" or recognition latency sense. At home or in a store, for example, greater familiarity or faster recognition of one item vs. others may be transformed into a preference indication by the child through the simple act of pointing. And even if the child can not only point to but can also name the cereal verbally, the process involved may still be one of imitative (in this case auditory or "echoic") image retrieval. This contrasts with a more adult-like symbolic information retrieval and information evaluation process.

The process distinction between child and adult modes of decision-making also has an important implication for preference or response measurement. Images are characterized as being "stimulus faithful" representations of the original stimulus or referent. Examples would be an image of a cereal carton in the visual mode or of a cereal brand name in the auditory mode. Symbolic representations, in contrast, bear only an arbitrary relationship to the original stimulus or referent. Answering a verbally stated preference question or considering an equally abstract numerical rating scale are tasks which require symbolic transformation ability. The child may only have stored the attribute data in a pre-symbolic, imagery form. Thus he or she may not be able to indicate preference in a verbal questioning procedure or on a rating scale, yet a preference may clearly exist. Alternatively, preferences may differ as a function of the particular mode of response used in measurement, especially with children. This possibility has not been explored in children's consumer decision research.

An adequate exploration of the child's visually operative data base of cereal attribute information requires an appropriate response mode and a reasonably wide range of output opportunity. This was achieved in the present study by asking children to draw cereals, free-hand, from memory. They were asked to draw a generic cereal box, i.e., just "a cereal"; the back of a cereal box (a primary location for promotional material); their favorite cereal; the cereal their parents eat; and a cereal that is "healthy and good for you when you are growing up." These five tasks provided multiple opportunities to assess visually recallable information concerning brand identification, nutrition, sweetness or sugar content, premium offers, and other salient attributes of cereals.

The children were then asked to describe, verbally: their favorite cereal; parents' cereal; and a healthy cereal. For fairly obvious reasons of abstraction difficulty the children were not asked to describe a generic cereal. For the three more concrete tasks, verbal attribute emphasis could thus be compared with visual attribute emphasis as a test of the general research hypothesis of uniquely stored visual attribute knowledge.

## Method

**Design.** Because it was essential that the visual information measures not be contaminated by prior verbal recall, the drawing tasks were administered first. Children received the drawing task instructions in the same order: generic, back of box, favorite, parents' and healthy. They then received the verbal questions in the same order: favorite, parents', and healthy.

The attribute emphasis variables were scored by coding the drawings and the verbal protocols on a simple attribute present/attribute absent basis. Coding was done by one experimenter (J.T.) and checked by J.R. The dichotomous judgment procedure produced very few disagreements.

**Subjects.** Participants in the study were 60 boys selected from two Philadelphia area schools of similar middle class enrollment. Three grade levels were included--1st, 3rd and 5th--with 20 Ss per grade.

**Apparatus and materials.** The children were provided with 6 or 7 sheets of plain white 8 1/2 x 11" paper. Each child was given an identical box of 10 new felt-tipped pens for the drawings. Each box was used three or four times, but carefully checked for new pen condition before repeated use. Colors provided in the box (Bic "Bananas") were red, green, blue, yellow, light blue, orange, black, purple, brown and pink.

**Procedure.** The study was administered to groups of 5 or 6 children from each class, using a large room in which the children were spaced to prevent sight of one another's drawings, with their work desks facing the outer walls. Two experimenters were on hand to read the instructions, check the children's general progress (though they had unlimited time if needed) and administer the verbal questions.

Instructions for the cereal drawing task were as follows:

- 1a. Draw a cereal box using these color pens.
  - b. (Probe:) Have you drawn in everything you can remember about the cereal box?
- 2a. Would you draw me the back of the cereal box now.
  - b. (Probe:) Have you drawn in everything you can remember about the back of the cereal box?
- 3a. Draw your favorite cereal.
  - b. (Probe:) Have you drawn in everything you can remember about your favorite cereal?
- 4a. Draw the cereal your parents eat.
  - b. (Probe:) Have you drawn in everything you can remember about the cereal your parents eat?
- 5a. Draw a cereal that is healthy and good for you when you are growing up.
  - b. (Probe:) Have you drawn in everything you can remember about a cereal that is healthy and good for you when you are growing up?

While drawing, children who asked how to spell words (e.g., the name of the cereal) were told, "Spell it as best you can." Children who asked how to draw the boxes, how big, etc., were told, "Any way you want to draw it."

Questions for the verbal attribute measures were as follows:

- 1a. What is your favorite cereal?
- b. What kind of cereal is it? (Probe:) What else

can you tell me about your favorite cereal?

- 2a. What is the cereal your parents eat?
  - b. What kind of cereal is it? (Probe:) What else can you tell me about the cereal your parents eat?
- 3a. What is a cereal that is healthy and good for you when you are growing up?
  - b. What kind of cereal is it? (Probe:) What else can you tell me about a cereal that is healthy and good for you when you are growing up?

The verbal interviews were administered individually. The children were then returned to the classroom after being asked not to tell others about the questions and drawings until tomorrow (it was a one day administration per grade).

## Results and Discussion

### The General Hypothesis

Children clearly do have strong visual imagery of cereals. The drawings themselves were quite remarkable for their overall conceptual execution (in a general "schema" sense) and for their detailed inclusion of attribute data. The imagery evident in the drawings was strong and vivid enough to indicate that children could well retrieve and utilize these iconic representations in decision making. Some most interesting evidence that visual imagery does include preference information was obtained in the fact that the generic cereals (i.e., "Draw a cereal box") were branded spontaneously by a majority of the children (Table 1). At first grade 67% of the children branded their generic concept drawings; this increased to 100% at third and fifth grades.

### Dual Code Usage

Brand names are an attribute that can be represented in both the visual and verbal output modes. The extent to which children used both modes is also shown in Table 1. Brand names were "mentioned" both visually and verbally (i.e., printed or partially printed in correct visual location on the drawing<sup>2</sup> and mentioned in response to the verbal questions) by only about 60% of first graders. However, this double code retrieval rose to 84% at third grade and 93% by fifth grade. The increase in double code retrieval held for the three tasks for which visual and verbal responses were required: favorite, parents' and healthy (the percentages just referred to are averages of these three tasks from Table 1).

### Brand Matching

Of critical interest in terms of the general research hypothesis predicting unique usage of visual and verbal codes are the rows in Table 1 showing the match between visually and verbally supplied brand names.

On average, the brands nominated in the two codes (within child) matched only 52% of the time for first graders with, again, a marked increase to 59% at third grade and 77% by fifth grade.

Looked at another way, the uniqueness contributed by the

<sup>2</sup> Some of the attempts by younger children to draw in the brand name were not fully decipherable; in these instances, the apparent brand identification was cross-checked with correctness of color-of-carton cues (see section on color matching).

TABLE 1

## BRAND NAME EMPHASIS IN VISUAL AND VERBAL RECALL

Cereal task criterion	Grade 1	Grade 3	Grade 5
<u>Generic</u>			
• Visual	.67	1.00	1.00
<u>Back of box</u>			
• Visual	.67	1.00	1.00
<u>Favorite</u>			
• Visual	.58	.71	.95
• Verbal]	.95	.94	.95
Both codes	.58	.71	.95
Brand match	.42	.53	.74
Visual uniqueness	.27	.44	.22
<u>Parents'</u>			
• Visual	.58	1.00	1.00
• Verbal	.92	1.00	.94
Both codes	.58	1.00	.94
Brand match	.50	.58	.76
Visual uniqueness	.14	.42	.24
<u>Healthy</u>			
• Visual	.68	1.00	1.00
• Verbal	.84	.80	.90
Both codes	.68	.80	.90
Brand match	.63	.65	.80
Visual uniqueness	.08	.35	.20

visual data--brand name salience that would have been missed by the verbal procedure alone--was 16% at first grade (though remember that fewer brand names were included by the 6 to 7 year olds), rising to a very substantial 40% at third grade, and 22% at fifth grade.

The finding that uniqueness was highest at third grade, coupled with the incidence of matching at each age level, suggests a tentative set of hypotheses about children's visual and verbal code usage in their "internal search" for brand information. The 6 to 7 year-old first graders seem to be characterized by "differential focus" on the two codes, with somewhat unclear focus on either. The 8 to 9 year-old third graders seem to be characterized by a period of clear focus accompanied by "maximal discrimination and perhaps conflict" between the two codes. The 10 to 11-year-old fifth graders show a "coincidence of codes." The older children may be becoming more adult-like in their cognitions in that their visual and verbal referents are more likely to coincide.

Support for these contents comes from quite a different set of dimensional results: the colors which the children employed in their cereal box drawings (Table 2).

The top row of each sub-table provides a measure of the vividness of children's imagery as indexed by the (average) number of different colors the children used. Here it can be seen that for four of the five tasks vividness reached a peak at third grade. Taken over all tasks, first graders selected an average of 3.24 different colors from the available set of ten; third graders, 4.52; and fifth graders, 4.24. This curvilinear trend in vividness tends to support a model of maximal dual code discrimination at about age 8 or 9.

TABLE 2

VISUAL COLOR USAGE (VIVIDNESS), COLOR MATCHING, AND CORRECTED "HIT RATE"<sup>a</sup>

Cereal task criterion	Grade 1	Grade 3	Grade 5
<u>Generic</u>			
Vividness	4.17	5.95	5.10
Color match	1.50	3.15	3.50
Corrected hit rate	.36	.53	.69
<u>Back of box</u>			
Vividness	2.83	4.25	3.70
Color match	.94	2.15	2.50
Corrected hit rate	.33	.51	.68
<u>Favorite</u>			
Vividness	3.21	4.82	4.63
Color match	1.53	3.12	3.05
Corrected hit rate	.48	.65	.66
<u>Parents'</u>			
Vividness	3.17	3.67	4.00
Color match	1.25	2.08	2.41
Corrected hit rate	.39	.57	.60
<u>Healthy</u>			
Vividness	2.84	3.90	3.75
Color match	1.32	1.40	1.95
Corrected hit rate	.46	.36	.52

<sup>a</sup>Vividness = mean number of colors used; Color matching = mean number of correct matches with actual carton colors; Corrected hit rate = ratio of color matches to number of colors used.

## Color Matching

Color matching gauges the extent to which the colors selected by the child matched the colors on the actual cereal box. Matching was ascertained by purchasing all of the cereals nominated by the children and comparing them with the individual drawings. A "match" was scored only if both the color and the location were correct.

Color matching provides a second type of evidence for the original research contention that visual retrieval can play a guiding role in preference since the color and its carton location would presumably have to be retrieved from visual memory. Matching data for colors are shown in the second row of the sub-tables in Table 2. Also, to adjust for any chance matches due to the sheer number of colors each child used, a corrected hit rate was computed. This is shown in the summary row of the sub-tables.

Even the youngest children were able to match correctly at a rather high rate. Again remembering that only about two-thirds of the drawings were brand-identified at this age, the first graders color matched correctly 40% of the time over the five tasks. The color match rate then rose linearly to 52% at third grade and 63% by fifth grade. It is interesting to note that color matching was highest for "favorite" cereals at the two younger grades (it may have asymptoted at fifth).

Taken overall, the demonstration of color matching of cereal packages in the visual mode yields powerful evidence that visual imagery is alone sufficient to influence preference. All that would be required is visual recognition by color. It might be noted in this connection that if a child were asked to color match verbally, i.e., to describe from memory the colors on a cereal carton, he would almost certainly have to "read" the description from a visually stored image.

TABLE 3  
ATTRIBUTE EMPHASIS

Attribute (by task criterion)	Grade 1	Grade 3	Grade 5
<u>Sugar sweetness<sup>a</sup></u>			
Favorite - verbal	.42	.41	.26
Parents' - verbal	.17	--	.12
Healthy - verbal	.05	--	.20
<u>Nutrition<sup>a</sup></u>			
Favorite - verbal	.10	.18	.58
Parents' - verbal	.25	.42	.53
Healthy - verbal	.37	.50	.60
<u>Premium<sup>b</sup></u>			
Generic - visual	--	.15	.45
Back of box - visual	.39	.55	.75

<sup>a</sup>Visual responses could not be accurately coded.

<sup>b</sup>Verbal task not administered for generic or back of box criteria.

#### Attribute Emphasis

Three primary cereal attributes were examined in the present study: sugar (or sweetness) emphasis, nutrition emphasis, and the presence of premiums as a means of promoting the cereal.

Sugar emphasis. Sugar or sweet flavor emphasis is difficult to represent visually; and also, for two or three years now, "sweet" claims have rarely been emphasized explicitly in cereal packaging or in television (visual) advertising. It is not surprising, therefore, that apart from occasional flavor specification in the drawings, visually represented sugar attributes did not appear. Verbal attribute emphasis for sugar and sweetness, however, was most interesting (Table 3; as described earlier, verbal questions were not asked for the generic or the back of box tasks). For favorite cereals, sugar or sweetness attributes were mentioned by 42% of the first graders and 41% of the third graders, but by only 26% of the fifth graders. For parents' cereals, sugar mentions were really low: 17% or less at all grades. For healthy cereals only 5% of first graders and none of the third graders mentioned sugar or sweetness; the fact that the sugar figure then rose at fifth grade, to 20%, could indicate

a very sophisticated response by these older children! Overall, the pattern is an encouraging one from an "educational" standpoint, although this was, as noted, a generally well educated research sample.

Nutrition emphasis. Nutritional attributes (mentioned as "vitamins," "natural," etc.) might be regarded as a counterpart to sweetness. These data are also shown in Table 3. Once again the results show an encouraging trend. As might be expected, the highest nutrition mentions were for the healthy cereals: 37% at first grade, 50% at third, and 60% at fifth. Also of interest was the sharp increase at fifth grade where 58% of the children also mentioned nutrition as an attribute of their favorite cereal.

Premium emphasis. Premiums are clearly a salient feature in children's visual memory of the generic concept "cereal." This emphasis increased with age. Figures for the back-of-box drawing task, where premiums are typically displayed, were 39% at first grade, 55% at third grade, and 75% at fifth. Front-of-box premium inclusion was somewhat lower, as might be expected, but still showed the increase: 0% at first grade, 15% at third grade, and then 45% at fifth grade. The latter finding is interesting because it may reflect increasing awareness and comprehension of claims about premiums (a symbolic verbal element) rather than simply iconic representation of the premium in visual memory.

The high incidence of premium inclusion in children's generic conceptions of cereals and the linear increase with age suggest that premiums do indeed become perceived as an attribute of cereals. Administration of a back-of-box drawing task for children's favorite cereals may have revealed even greater salience of premium information. This finding is of considerable interest in the light of recent controversy over whether children comprehend premiums as part of the product or as a separate choice consideration. A children's cereal without a premium may be lacking an expected attribute.

#### Conclusions

The intent of this study was to demonstrate that children have a rich data base of information about products (such as cereals) which is stored in visual rather than verbal memory and which can alone influence their preference process.

It was shown that retrieval of visual information can lead to quite a different indication of children's brand preferences. Visual and verbal nominations of the child's favorite cereal, for example, only corresponded in 42% of instances at age 6 to 7, 53% at ages 8 to 9, and 74% at ages 10 to 11. Visual representation therefore contains a very substantial amount of information (ranging from 8% to 44% of brand preferences) that would be missed by traditional verbal or numeric measures of children's preferences.

It was further shown that visual representations of cereals are sufficiently strong or vivid to be utilized in children's preference decisions without symbolic verbal assistance (e.g., in a supermarket situation, visual recognition then pointing). This was evidenced by the fact that even the youngest children could correctly color-match the details of actual boxes with 40% accuracy (48% for their favorite brand), a figure which increased to 63% for the older children. Colors must almost certainly be retrieved from purely visual memory images of the cereals.

Some attributes such as sweetness or nutrition claims are, by nature, difficult to encode visually. However, it was found that another dominant attribute of cereals--promotional premiums--can be and are coded visually. When asked to draw the back of a cereal box 39% of first graders included a premium, 55% of third graders did so, and 75% of fifth graders. Unfortunately, the back-of-box task could not be included for every drawing, so the extent to which premiums are salient to the child's favorite cereal concept was not fully ascertained. Premiums did, however, receive very low verbal recall, suggesting again the uniqueness of visual retrieval.

The visual and verbal data obtained in the study also showed highly consistent trends (generally linear increases) with age. One exception was vividness of visual imagery, which appeared to peak in curvilinear fashion at third grade, an age at which the discrepancy between visually and verbally produced preferences also happened to be at a maximum. This suggests at least one innovative direction for further research. The challenge from a research design standpoint and the fascinating output of visual memory measurement should be enough to stimulate many others.

To reiterate the issue with which this research on visual memory began: consumer preferences may be based on information other than that represented in the form of verbal (or verbalizable) beliefs. Consumers may utilize information stored uniquely in visual memory. This data base may not be amenable to retrieval via the typical verbal measurement procedures such as those employed, for example, in conjunction with "Fishbein-type models."

The present study focused on visual and verbal memory in children's consumer behavior. Visual memory may be more important than is commonly realized in adult decision making. For example, all of us from time to time consult our "mental maps" of shopping environments and store layouts to facilitate our consumer decision processes. Once we are in the store, in a supermarket for instance, we may then choose to "breeze through," making many of our brand selections in a rather uninvolved, rather automatic way. In doing so we may be relying on momentary evocation of attitudes or preference indications from a data base located primarily in visual memory.

#### Appendix

##### Tabular Base Levels

Tables are based on N = 20 per grade level with the following exceptions: 2 first grade children could not conceptualize a generic cereal even after two or three instruction repetitions; 1 first grader, 3 third graders and 1 fifth grader did not have a particular favorite cereal; the parents of 8 first graders, 8 third graders and 3 fifth graders did not eat cereal, thus making the parents' cereal task irrelevant for them; and 1 first grader could not conceptualize a healthy cereal. Base levels for proportion data in the tables were adjusted accordingly.

#### References

Bobby J. Calder, Thomas S. Robertson and John R. Rossiter, "Children's Consumer Information Processing," Communication Research, 2 (July, 1975), 305-316.

Allan Paivio, Imagery and Verbal Processes (New York: Holt, Rinehart & Winston, 1971).

Michael I. Posner, Cognition: An Introduction (Glenview, Illinois: Scott, Foresman, 1973).

## TELEVISION COMMERCIALS

### AS SOCIALIZING

#### AGENTS

Ann H. Beuf, The University of Pennsylvania

#### Abstract

The paper addresses the question of the means by which television commercials, as well as performing their manifest function of promoting a product, also act as socializing agents in American Society. There are essentially three parts to this consideration. First, some theoretical background from sociological work in the area is provided, dealing with secularization and the assumption of value-transmitting tasks by secular institutions. The manner in which television-commercials-as-socializers may be viewed as a modification of the concept of American way of life is discussed. Next (1) the messages of commercials and (2) the devices of presentation of commercials are examined with regard to how they function to convey certain primary American values to children. Finally, the above observations are approached from three different perspectives on socialization. It is suggested that television with its extensive use among American children has the potential to function as a common conscience which introduces and reinforces American values at a high level of generality and that research in this area should take this dimension of the experience of commercial viewing into account in assessing the quality of children's commercials.

#### Introduction

It is quite clear to all of us that changes have taken place in the socialization of children during the past several decades. The increasing differentiation of school from family has expedited changes in this area; many functions once performed by parents are now part of the teacher's role. Another process has occurred concurrently with this process of differentiation of family and education with regard to the 'secondary socializers' in a child's life. In Athenian Greece (Payne, 1916) in Church-controlled Europe preceding the reformation, and in the Puritan era which followed it, religion functioned as a conveyor of important societal values (Payne, 1916). This function has diminished in an increasingly secular era. In this paper television, especially commercials will be considered as a new secondary socializer on the American scene, stepping in to perform some of the tasks of value definition and inculcation which are no longer so effectively performed by sacred institutions.

There has been much sociological speculation on the nature of the secularization of our society. Some have argued that Americans still relate to sacred institutions but do so for secular reasons such as the need for ethnic identification (Herberg, 1960); others (Cox, 1966) have stressed the growth of a new "people-centered" religion based on the principles of Autonomy, the dictates of conscience and the performance of good acts towards others within the 'secular city'. Another approach to the question of secularization has been the notion of a fusing of religion and patriotic rationatism. (Bellah, 1967).

Herberg also argues for the co-existence of formal religious institutions and another, secular system of ethics which he calls the American Way of Life. This system embodies Bellah's Civil Religion, but in addition includes other values not so directly related to patriotism: "there is to be found among Americans some sort of faith or belief or set of convictions, not generally defined as religion, but definitely operative as such in their lives in the sense of providing them with some fundamental context of normativity and meaning." (Herberg, 1960). Herberg goes on to note that it is the American way of life which supplies Americans with an "overarching sense of unity," and a framework in which "crucial values" of American life are couched. (Herberg, 1960) Some of the prominent features of this system are: love of fellow man, patriotic zeal, activism, individualism and achievement. The roots in Protestantism are immediately apparent -- in fact Herberg calls it "a sort of secularized Puritanism." (Herberg, 1960) This value system is disseminated by patriotic ceremonies in schools, public ceremonies (such as presidential inaugurations), and the celebration of holidays which commensurate events in the nation's history. (Bellah, 1967) (Herberg, 1960)

#### An Extension of the Theory

It seems necessary to consider one of the limitations of this approach which has special relevance for this paper. While the American Way of Life is a useful concept, it seems doubtful that it would be as widespread as Herberg claims it to be if dissemination were indeed dependent on schools, families and celebrations. Participation in and acceptance of these institutions is not so universal as to provide a really common set of experiences and a core of common values for the entire society. For a variety of reasons, many groups within American society are alienated from, ignorant of, or denied access to the very institutions which both Bellah and Herberg see as primary disseminators of a core value system. (Bellah, 1967) (Herberg, 1960)

For the poor, the minority person, involvement in such events as the fourth of July, Washington's Birthday and Thanksgiving (this last being a real bete noir of the American Indian community) is either absent or occasion for ambivalent affect. The present debates raging within several minority communities over whether or not they should participate in the Bicentennial celebration is frank evidence of the lack of a societywide acceptance of certain elements of the American way of life.

I believe that Herberg is essentially correct that secular vehicles are increasingly employed to inculcate social values, but I would extend his argument. I believe that many other institutions, among them the mass media, perform the function of value dissemination and are better able to do so in the face of political and social diversities and animosities than are traditional institutions.

The central contention of this paper will be that there is within our society a socializing agent which reaches large numbers of children, regardless of class or race, thus providing them with (a) a common set of experiences and (b) exposure to a common set of values while manifestly performing other tasks of entertainment or salesmanship. This secular vehicle is television and I shall argue that it invokes the same value clusters as those which are present in both the American way of life and the Protestant Ethic as delimited by Max Weber (Weber, 1958) Television as a Socializer.

In socializing the American child to a set of values both message content and mode of presentation are invoked. Some commercials, notably public service messages, overtly instruct a child in such values as honesty (anti-shoplifting) and neatness ("Don't be a litterbug!"). These messages, direct and to the point, represent an obvious attempt to give instruction in "correct" behavior. What we shall be concerned with here, are the more subtle cases, those in which human virtues are being peddled along with cereal and sneakers.

The values most frequently stressed in the television commercials I have observed are achievement-success, individualism, equality, patriotism, fairness, and equality, good health, a rather peripheral component of the American way of life is also frequently mentioned, especially in food commercials. All of these values are important elements of The American Way of Life as Herberg has delineated it. (Herberg, 1960) These values are stressed in a variety of ways, some quite direct, others rather subtle. I have noted five ways of introducing or reinforcing an American Way of Life value in children's commercials.

A. Direct instruction. Here the audience child is told what or what not to do. Examples: "Don't be a litterbug!" "Hug your Baby Tender Love. Love her."

B. Indirect instruction. Here the television actor child is told what or what not to do. Thus an adult figure tells a child to share his Nabisco cracker with his sister. In a Capt'n Crunch commercial when one child starts to help another with a riddle he is reprimanded "Don't tell him. Let him do it himself."

C. Modeling. This device simply approaches the viewer by showing good things happening to children because of something they do. Winners receive pats on the back -- a tacit reward for achievement. Girls who are "friendly" are shown drinking Kool-Aid with friends. Sometimes this modeling has to do with who the children are with. In an effort to promote Equality, many children's ads are now multi racial. In the television world children in these groups are always smiling and laughing.

D. Overvoice. The use of adult, usually male adult overvoices to tell about the product while children are seen using it, can be viewed as a cultural way of lending authority to the values stressed and also of reinforcing the obedience value in viewers.

E. Appeals to children's anxieties. Some commercials speak the fears and anxieties of children. Is one feeling small and vulnerable -- this toy you can use "all by yourself." Is a little girl jealous of her mother -- with Baby That-A-Way she can identify with her. Is junior in the throes of the Oedipus complex -- he can beat hell out of his Dad at Parcheesi. If he "plays his hardest."

F. Appeals to Pity. Here ill or poor persons are shown in a condition such as to stir up in the child the proper emotions of sympathy and generosity (which

are then commended by the male adult overvoice).

Almost every device is employed in the service of the American Way of Life. Some mechanisms seem to be favored for the inculcation of certain specific values. For example, I observed that individualism is very often conveyed through the mechanism of the overvoice addressing the viewer. "You should try this yourself." "You can build this unaided," while equality (perhaps because of its dubious standing in the value hierarchy right now) is never referred to verbally, but is conveyed through a portrayal of happy interracial groups. Patriotism is most frequently introduced through the juxtaposition of a common patriotic symbol with an object which has good connotations. Example: Grandpa Strohman singing to the tune of Yankee Doodle Dandy amid stars and stripes while enjoying "Some of mine delicious bread."

What becomes evident through prolonged viewing is that the American Way of Life is being sold by reference to toys and candy, just as toys and candy are being sold by references to the American Way of Life. So far, investigators have only dealt with the latter aspect of this two-fold phenomenon.

#### Socialization Theories

On a theoretical level, at least, we can interpret these observations from several perspectives in the area of socialization. (Empirical investigation will have to be carried out to find which of these, or which combination of them is actually at work.)

From a learning theory perspective we can see children as receiving direct instructions and indirect messages via models on the screen as they view television commercials. "Try doing this yourself first!"

The positive reinforcement received by children on the screen with whom the viewer identifies can also be employed as a socializing device (Bandura, 1967). Thus the smiles of approval which are directed at the winner of a game, reinforce aggressiveness and success as values. (not just the value of the table game!)

From a symbolic interactionist point of view, television and those who people it can come to be internalized in personality as both "specific others" and "generalized others." (Mead, 1934). Thus the expectations for social behavior expressed and demonstrated on television become a part of the child's value repertoire.

Finally, Freudians might claim that the special psychological anxieties and fears of children are manipulated by the symbolism of television commercials. Certainly doll commercials capitalize on the female child's wish for a baby, while instilling ideas of "proper mothering," and the frequency with which sons are portrayed beating Dad in game commercials does cater to Oedipal fantasies -- associated with aggressive game behavior as well as with the game itself.

#### Conclusion

It appears then, that television and in particular television commercials function as socializers of American children by selling values as well as products. The process involves seems to be many-faceted and probably involves the child in cognitive learning; internalization of the norms and values held by the "generalized other" such as it is presented in the world of television, and a tapping of deeper psychological concerns and anxieties. It has



been suggested that this mechanism has, along with such other institutions as the schools and the political-legal system, emerged as an important "secondary socializer," and that this has accompanied the diminished importance of religious institutions in this realm. The family probably remains the most important socializing agent for our children, but increased use of television (Schramm, 1968) and the long hours spent in its presence by most American children have given it a role in their lives which we should not ignore. That it reaches and holds the attention of a wide variety of children regardless of race or social class may make the watching of television and the learning from it the most common experiences of young Americans as a people. Regardless of how widely different may be the experience of the Eastern Urban Black youngster and the Anglo farmers child of the Mid-West, in all likelihood these two children, if they met could immediately exchange reminiscences about the adventures of Speed Racer, Bugs Bunny and the crew of the Starship Enterprise. They doubtless also could recite the lyrics of many a commercial and reel off the chemical contents of a variety of toothpastes and other toilet articles.

This idea of a shared experience was a primary ingredient for the development of a "common conscience" as it was developed by Emil Durkheim. (Durkheim, 1964) The common conscience was the source of religion and was reflected by it. The common consciences, or shared value systems of small cohesive groups were believed by Durkheim to have been weakened by Western Society's movement toward large, heterogeneous, industrial-urban societies. It appears, however, that for the coming generation of American children, the common conscience may be rising, Phoenix-like from its ashes to manifest itself in children who have been exposed to the same set of values by the mass media.

Children are not just exposed to a series of isolated values through this process. As Gross has noted (Gerbner and Gross, 1974) they are presented with an entire world view. Furthermore, it seems likely that they will buy this world view, precisely because the fact that it is being conveyed is a disguised or latent process. Robertson and Rossiter (Robertson and Rossiter, 1974) have indicated that the suggestions children resist are likely to be those which they recognize as having persuasive content. As the value messages are hidden dimensions of commercials, the child is unlikely to realize that he is being "sold" a value, and therefore unlikely to "turn-off" to it.

In light of these considerations it seems important for us to give serious thought to the nature of the television-world-view as it is conveyed by commercials. The values it applauds, the degree to which it represents either the static complacency of the status quo, or a liberating challenge to existing repressive social arrangements should be the subject of thorough and sensitive research in the fields of communications and consumer behavior.

#### References

Albert Bandura, "The Role of Modeling Processes in Personality Development," in Willard W. Hartys and Nancy L. Smothergill (Eds.), The Young Child: Reviews of Research (Washington, D.C.: National Association for the Education of Young Children), 1967.

Robert Bellah, "Civil Religion in America," Daedalus, 96 (Winter 1967) 1-22.

Harvey Cox, The Secular City (New York: The Macmillan

Company, 1966)

Emil Durkheim, The Division of Labor in Society (New York: The Free Press, 1964).

George Gerbner and Larry Gross, "Cultural Indicators: The Social Reality of Television Drama" (Unpublished paper, 1979).

Will Herberg, Protestant, Catholic, Jew (Garden City, New York: Anchor Books, 1960).

M. L. Kohn, "Social Class and Parent-Child Relationships," American Journal of Sociology, 68 (1963) 471-480.

George H. Mead, Mind, Self and Society, C. W. Morris (Ed.) (Chicago: University of Chicago Press, 1934).

George H. Payne, The Child in Human Progress (New York: G. P. Putnam's Sons, 1916).

Thomas Robertson and John R. Rossiter, "Children and Commercial Persuasion: An Attribution Theory Analysis," Journal of Consumer Research, 1 (June 1974), 13-20.

William Schramm, Jack Lyle and Edwin B. Parker, "Television in the Lives of our Children," in Derek Phillips (Ed.), Studies in American Society: I (New York: Thomas Y. Crowell Company, 1968).

Max Weber, The Protestant Ethic and The Spirit of Capitalism (New York: Charles Scribner's Sons, 1958).

THE DEVELOPMENT OF CONSUMER INFORMATION-PROCESSING SKILLS:

CONTRIBUTIONS FROM COGNITIVE DEVELOPMENT THEORY\*

Daniel B. Wackman, University of Minnesota  
Scott Ward, Harvard University  
and Marketing Science Institute

Abstract

The combined theoretical perspective we have been using in studying consumer socialization--cognitive development theory and an information processing model--are described. Some data from our prior research is presented to illustrate the usefulness of the perspective. Lastly, directions for further research with the combined perspective and its relevance for policy questions are discussed.

In recent years, researchers, practitioners, and policy makers in industry and government have evidenced considerable interest in the topic of "consumer socialization." Under this rubric, the specific interest is often in one of three areas. First, researchers have been concerned with television advertising practices affecting children, and related policy issues (e.g., Ward, 1972; Rubin, 1972; Shimp, Dryer and Divita, 1975). Second, researchers have been interested in marketing management issues involved in promoting to children (McNeal, 1964; James, 1971; Schiele, 1974). Third, recent attention has been devoted to the topic of consumer socialization itself. The topic is usually defined as having to do with the processes by which children acquire knowledge, skills and attitudes related to consumer behavior. Specific research attention in this area has focused on such topics as parent-child interaction (Ward and Wackman, 1972; Berey and Pollay, 1968), adolescent consumer learning (Moore and Stephens, 1975), and theoretical issues involved in consumer socializations processes (Ward, 1974; Calder, *et al.*, 1975). In short, the area is an appealing one for research, since it can target on immediate policy questions, management questions, and longer-term developmental issues.

The area of consumer socialization is an appealing one for other reasons, as well. It suggests some reasonably well-defined problems, and the investigator is led rather directly into a set of reasonably comprehensive theories which are relevant to the problems. This is in marked contrast to the reverse process which typically occurs in consumer behavior research, i.e., investigators borrow theories from behavioral science, and search for applications in "real" marketing problems.

The Cognitive Development Approach  
to Consumer Socialization

Consumer socialization can be viewed as a particular aspect of more general socialization processes. That is, the interest is in how children acquire knowledge, skills and attitudes relevant to marketplace activities, as distinct from other activities. Two theoret-

tical approaches have been particularly useful to consumer socialization researchers to this point: learning theory approaches (e.g., Goldberg and Gorn, 1974; Calder, Robertson and Rossiter, 1975), and cognitive development theory.

Learning theory is a broad term, and its relevance to socialization research would seem to stem from three particular approaches: neo-Hullian, neo-Skinnerian, and social learning theory (see Zigler and Child, 1969, pp. 465-468). The latter has been cited as holding much promise for consumer socialization research, among learning theory approaches (Calder, Robertson and Rossiter, 1975; Atkin, 1975).

We have found cognitive development theory to be most useful for three primary reasons. First, cognitive development theory, stemming primarily from Piaget (1928, 1952 and 1954), is directly relevant to issues relating to children's information processing of consumer stimuli. That is, developmental concepts are useful in understanding and predicting children's selection, evaluation and use of information in consumption decisions. Second, cognitive developmental theories are directly relevant to policy issues, and we believe that one criterion for theory is its utility for employing concepts and generating research which is relevant to practice. Developmental theory postulates age-related changes in cognitive abilities which clearly bear on children's information processing abilities and strategies; consequently, the theory is useful in guiding policy decisions concerning what are and what are not appropriate marketing stimuli for children. Finally, widespread application of cognitive development theory has resulted in a high degree of face validity for research results. In our own research, for example (Ward and Wackman, 1973) we were able to predict several dimensions of children's reactions to commercials on the basis of cognitive developmental concepts, and used several methods to test individual hypotheses concerning stage related differences in these reactions. As Calder, *et al.*, (1975) accurately point out, cognitive development theory specifies "limitations" on cognitive processing abilities among children, and the theory cannot be extended too far in explaining particularistic phenomena (e.g., children's selection of bits of information from specific kinds of commercials). It is misleading to conclude, however, that developmental theory lacks explanatory or predictive power. In any case, the issue really concerns what we expect from theory. Our position is that many crucial issues concerning children's developing abilities to process consumer-related stimuli are well accounted for by cognitive developmental concepts.

Overview of Cognitive Developmental Concepts --

As with all cognitive developmental theories, Piaget's theory posits a cognitive representational or coding process intervening between a stimulus and a child's

\* To appear in *Advances in Consumer Research*, Vol. III, (Proceedings, Association for Consumer Research Annual Meeting, Cincinnati, Ohio, October 1975).

response. These cognitive representations are organized into cognitive structures, or intergrated patterns of thought and behavior which change with age. Kohlberg (1971) refers to these cognitive structures as "rules used for processing information and connecting experienced events." The qualitative changes in thought which children undergo are conceptualized as "stages," which are considered as (1) distinct, qualitative differences in children's modes of thinking or problem solving; (2) invariant so that environmental facts may influence the pace of development, but not the sequence; (3) structured wholes, so that a child at a given stage will show thinking typical of his stage in diverse situations; and (4) not in conflict with one another--each successive stage encompasses lower stages into higher levels of organization.

Piaget posits four developmental stages: (1) sensorimotor (birth to two years); (2) preoperational (two to seven years); (3) concrete operational (seven to eleven years); (4) formal operational (eleven through adulthood). The age ranges are approximate, but the stage-age relationship has been well established (Favell, 1963).

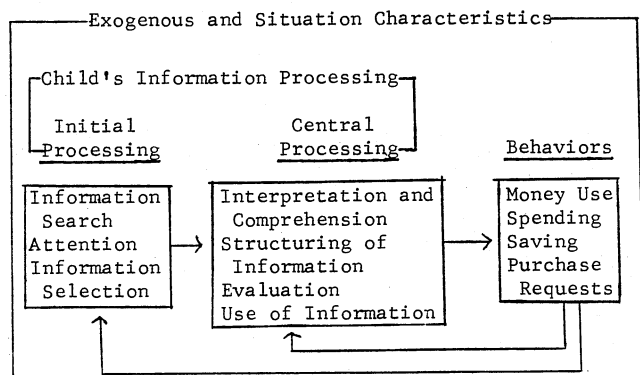
Of particular interest in current research are the differences between preoperational and concrete operational thinking since children in these two stages are the focal age group in current policy controversies. Preoperational thinking is characterized by developing symbolic abilities (such as language and mental imagery), but the child at this stage is very much "perceptually bound," i.e., reality is understood primarily in terms of the immediate perceptual environment. By the concrete operational stage, the child has developed conceptual skills which enable him to effectively mediate perceptual activity, but only when dealing with concrete objects. Children at the concrete operational stage can focus on more dimensions of a situation than can younger, preoperational children, who tend to focus on one dimension (the concept of "centration"). Other cognitive abilities differentiate preoperational and concrete operational developmental stages too, but our research has primarily focused on this perceptual boundness dimension.

#### Using Developmental Concepts in Consumer Socialization Research

Developmental concepts are useful in understanding differences in children's selection of information, and evaluation and use of information, in making consumer-related decisions. These processes are portrayed in Figure 1, as "initial" and "central" processing.

FIGURE 1

#### Research Model of Children's Consumer Behavior



Data relevant to children's "initial processing" of information are seen in studies by Ward and Wackman, 1973; Wartella and Ettema, 1974, and Ward, Wackman and Wartella, 1976). In these studies, the basic proposition was that younger (preoperational children) would exhibit greater responsiveness to perceptual characteristics of commercials. The expectation was supported, as indexed by attention behavior, and by post-exposure verbal data, in which younger children were most likely to recall and describe commercials in terms of salient visual stimuli.

It was also found that children exhibit qualitative differences in recall of commercials, as well as differences in the amount of material recalled. Data in Table 1, for example, show clear developmental differences in the degree of organization of recalled information, as predicted by development theory.

TABLE 1  
Recall of Favorite Commercial

	Grade			
	K	3	6	Total
Recall Unidimensional (Descriptive)	31%	12%	9%	17%
Recall Multidimensional and Random	50%	36%	27%	37%
Recall Multidimensional and Coherent	17%	48%	52%	40%
Recall Multidimensional, Coherent, and with Selling Message	2%	4%	12%	6%
	100%	100%	100%	100%
N =	(166)	(187)	(193)	(546)
	$\chi^2 = 93.40$ (6 d.f.) $p < .001$			

Another aspect of initial information processing focuses on the types of information children select in order to evaluate a product. Developmental theory would lead one to predict that preoperational children would be more likely to use perceptual (physical) attributes as a primary basis of evaluation, compared to older, concrete operational children, who should use more attributes, particularly "abstract" ones, in product evaluations. Data in Table 2 support this prediction: when asked about the kind of information they would want if they were buying a television set, most children requested information about physical attributes (how big is the screen?), but older children were considerably more likely than younger children to also request information about performance, price and functional (e.g., how long will it last) attributes.

TABLE 2  
Selection of Information About a  
Television Set (% Mentioning at Least Once)

Attributes	Grade				$\chi^2$ (2 d.f.)
	K	3	6	Total	
Physical	84%	87%	71%	80%	19.48
Performance	36%	49%	57%	48%	16.84
Price	15%	31%	41%	30%	27.99
Functional	13%	26%	43%	28%	38.21
N =	(157)	(199)	(202)	(558)	

As an aspect of "central processing," (see Figure 1), developmental theory was used to predict age-related differences in the number and kind of attributes children use to compare brands. Data in Table 3 show the mean numbers of attributes children used to compare two brands in three product groups (peanut butter, milk additives and toothpaste). All three age groups are similar in mentioning perceptual attributes of the products, but older children are more likely to mention ingredients and functional attributes which are more conceptual in nature (e.g., "the way the milk additive mixes with water" or "the toothpaste keeps you from getting cavities").

TABLE 3

Mean Number of Attributes Per Product  
Used to Compare Brands

	<u>Perceptual Attributes</u>	<u>Ingredients</u>	<u>Functional</u>	<u>Total Information</u>
K	.88	.17	.45	1.50
3	.93	.26	.66	1.85
6	.93	.35	.70	1.98

We should point out that the results reported here are not particularly surprising: older children use more, and more "sophisticated" kinds of information in their developing consumer information processing skills. The point is that the results are predicted from cognitive development theory.

#### Future Directions For Research

Up to this point, our research has focused on the content of children's consumer information--the kinds of information children select and use relative to consumption. The next step in our research will be to shift our focus to the processes children use in selecting, storing, and handling consumption related information. Two major areas of interest will receive our attention initially: representational processes and information use processes. These correspond quite closely to the distinction between initial processing and central processing depicted in Figure 1.

Representational processes refer to the processes by which children select information and store it in memory. Our research strategy to the present has been to ask children open-ended questions, thereby trying to tap verbally what children have stored. But, as Rossiter (1975) properly points out, alternative representations (e.g., visual) are possible and, indeed, they are likely, as his data show. Of particular interest in the area of representational processes will be examination of questions regarding how children store product-related information.

Information use processes refer to the processes by which children utilize information in making decisions. Recent consumer research has focused on a variety of strategies adults can--and do--use in choosing brands, e.g., various compensatory models, lexicographic models, conjunctive models, etc. To what extent children use simplified versions of these models--or entirely different ones--will be of central interest in our research. And obviously how a child represents information will have a major impact on the strategies he can adopt in making consumption decisions.

In terms of guiding this future research, we expect the

cognitive development literature to be just as useful as we have found it in our previous research. Specifically, we expect it to be useful in providing concepts descriptive of specific cognitive operations that can stipulate necessary conditions for utilizing different representational and information use processes.

For example, in focusing on the question of how children store information regarding various brands within a product category, we can ask what are the cognitive skills necessary to build an object/attribute matrix. We can stipulate a sequence of cognitive operations and the resultant representation that can be made when each operation is added:

1. Ability to see an object as separate from properties of the object (i.e., attributes)--having this ability makes it possible to construct a one object, one attribute matrix.
2. Ability to classify objects according to whether they possess an attribute or not and ability to compare two objects on an attribute dimension having these abilities makes it possible to construct a two object, one attribute matrix.
3. Ability to order a series of objects on an attribute dimension--having this ability makes it possible to construct a multiple object, one attribute matrix.
4. Ability to utilize more than one dimension--having this ability makes it possible to construct a multiple object, multiple attribute matrix.\*

In short, specification of the operations necessary to building a full matrix and the kinds of matrices implied by having only a partial set enables us to conceptualize--and then test--the kinds of representation children of various ages (and skill levels) are likely to have. Each of these cognitive operations, and the sequence of their development are discussed in cognitive development theory.

Similarly, different types of logical operations identified by cognitive development theorists will help in specifying different kinds of information use processes children might utilize in making consumption decisions. Cognitive developmental concepts that may be candidates for suggesting necessary conditions for utilizing various information use strategies include seriation, inferential ability, and transitive reasoning, etc. In general, it is to be expected that the child's ability or inability to perform various logical operations will (a) limit or expand the number and types of decision strategies a child can use, and (b) reduce or increase his flexibility in shifting strategies.

Besides suggesting a variety of interesting research questions, we believe the combination of cognitive development theory with an information processing model is highly relevant to many policy questions too. For example, take the premium advertising issue. An important question concerns whether advertising of a premium tends to sell the premium itself to the child, or whether it sells the primary product which has "premium" as one of its attributes. In our conceptual terms, this becomes a question of the child's representation of the information about the product and premium. At least two alternatives are clearly possible. (1) The product,

\* It should be noted that these cognitive operations assume that the information may be stored in various codes--verbal, visual and other types of representations are possible.

such as a cereal, and the premium are seen as two separate objects. (2) The premium is seen as an attribute of the primary product, perhaps a very important one, but nevertheless clearly a part of the product.

In Rossiter's (1975) research reported in this volume, a large percent of the children drew a premium on a cereal box when asked to draw a cereal box. He interprets this as indicating that children see premiums as an attribute of cereal. However, when he asked the children to verbally describe their favorite cereal, few made any reference to premiums; instead, they mentioned taste, sweetness, etc. This suggests the possibility that the subjects in Rossiter's study were responding to two different objects. On the one hand, they were asked to draw a cereal box, and for the children, premium is an important attribute of the object "cereal box." On the other hand, they were asked to describe their favorite cereal, and for them, premium is not an important attribute of the object "cereal." If this alternative interpretation is correct, then it would appear that most children are representing premium and cereal as two different objects. Of course, this would increase the likelihood that advertising of a premium is indeed selling the premium to the child, rather than the primary product. In any case, the example clearly indicates the relevance of questions raised when an information processing perspective is utilized.

To summarize our position briefly, we believe our prior research has shown that the combination of an information processing perspective and cognitive development theory is an extremely useful conceptual marriage in guiding our research and in providing explanations for our results. As we continue to pursue consumer socialization research--shifting our attention to questions regarding the processes children utilize in representing and using information--we are convinced it will continue to be extremely useful.

#### References

- Atkin, C.K., "Children's Social Learning From Television Advertising: Research Evidence on Observational Modeling of Product Consumption," ACR Proceedings, Vol. III, 1975.
- Banks, S. "Public Policy on Ads to Children," Journal of Advertising Research, Vol. 15, No. 4, August 1975, pp. 7-15.
- Berey, L.A. and R. W. Pollay. "Influence Role of the Child in Family Decision Making," Journal of Marketing Research, 5, February 1968, pp. 70-72.
- Calder, B.J., T.S. Robertson and J.R. Rossiter. "Children's Consumer Information Processing," in M.L. Ray and S. Ward, (Eds.), Communication Research special issue on Consumer Information Processing, Vol. 2, No. 3, July 1975, pp. 307-317.
- Flavell, J.H. The Development Psychology of Jean Piaget. New York: D. Van Nostrand Company, 1963.
- Goldberg, M.E., and G.J. Gorn. "Children's Reactions to Television Advertising: An Experimental Approach," Journal of Consumer Research, Vol. 1, No. 2, September 1974, pp. 69-75.
- Hill, R. and Joan Aldous. "Socialization for Marriage and Parenthood," in D. Goslin (Ed.), Handbook of Socialization Theory and Research. Chicago: Rand McNally and Company, 1971, pp. 951-983.
- James, L. Youth, Media and Advertising. Austin: University of Texas Bureau of Business Research, 1971.
- Kohlberg, L. "The Cognitive Developmental Approach to Socialization," in D. Goslin, (Ed.), Handbook of Socialization Theory and Research. Chicago: Rand McNally and Company, 1971.
- McNeal, J.U. Children as Consumers. Austin: University of Texas Bureau of Business Research, 1964.
- Moore, W.E. "Occupational Socialization," in D. Goslin, (Ed.), Handbook of Socialization Theory and Research. Chicago: Rand McNally and Company, 1971, pp. 861-885.
- Moore, R.L. and L.F. Stephens. "Some Communication and Demographic Determinants of Adolescent Consumer Learning," Journal of Consumer Research, Vol. 2, No. 2, September 1975, pp. 80-93.
- Piaget, J. The Child's Conception of the World. New York: Hartcourt, Brace, 1928.
- Piaget, J. The Origins of Intelligence in Children. New York: International Universities Press, 1952.
- Piaget, J. The Construction of Reality in the Child. New York: Basic Books, 1954.
- Rossiter, J.R., "Visual and Verbal Memory in Children's Product Information Utilization," ACR Proceedings, Vol. III, 1975.
- Rossiter, J.R. and T.S. Robertson. "Children's Television Viewing: An Examination of Parent-Child Consensus." (in press) Sociometry, 1975.
- Rubin, R.S. "An Exploratory Investigation of Children's Responses to Commercial Content of Television Advertising in Relation to Their Stage of Cognitive Development," unpublished doctoral dissertation, University of Massachusetts, 1972.
- Schiele, G.W. "How to Reach the Young Consumer." Harvard Business Review, Vol. 52, March/April 1974, pp. 77-86.
- Shimp, T.A., R.F. Dryer and S.F. Divita. "Advertising of Children's Premiums on Television: An Experimental Evaluation of the FTC's Proposed Guide," unpublished paper, School of Government and Business, George Washington University, 1975.
- Ward, S. "Consumer Socialization," Journal of Consumer Research, Vol. 1, No. 2, September 1974, pp. 1-15.
- Ward, S. "Kids' TV-Marketers on Hot Seat," Harvard Business Review, Vol. 50, No. 4, July/August 1972, pp. 16-37.
- Ward, S. and D.B. Wackman. "Children's Information Processing of Television Advertising," in P. Clarke, (Ed.), New Models for Mass Communication Research, Vol. II, Sage Annual Reviews of Communication Research. Beverly Hills: Sage Publications, Inc., 1973, pp. 119-146.
- Ward, S. and D.B. Wackman. "Children's Purchase Influence Attempts and Parental Yielding," Journal of Marketing Research, 5, February 1968, pp. 70-72.
- Ward, S., D.B. Wackman and E. Wartella. Children Learning to Buy: The Development of Consumer Information Processing Skills. (In press), Beverly Hills: Sage Publishing Company, 1976.
- Wartella, E. and J. Ettema. "A Cognitive Developmental Study of Children's Attention to Television Commercials," Communication Research, 1, January 1974, pp. 44-69.

Zigler, E. and I.L. Child. "Socialization," in G. Lindzey and E. Aronson, (Eds.), The Handbook of Social Psychology, second edition; Vol. 2 The Individual in a Social Context. Reading, Massachusetts: Addison-Wesley Publishing Company, 1969.

COGNITIVE RESPONSE TO ADVERTISING:  
THE RELATION OF CHILD TO ADULT MODELS

Bobby J. Calder, Northwestern University  
Thomas Robertson, University of Pennsylvania  
John Rossiter, University of Pennsylvania

Abstract

It is argued that child models of cognitive response to advertising are not solely constrained adult models. Nor is the relationship between child and adult models completely explained by Piaget's stage construct. Differences are discussed in terms of the child's mediational representation, organization, and evaluation of advertising information.

An important theoretical issue in investigating the cognitive response of children to advertising concerns the relationship between child and adult models of cognitive processes. It may be argued that any child model presupposes a well-developed model for adults. A model for children may be viewed as a constrained version of the adult model. The child model simply includes limits in the information-processing abilities of children. The problem with this argument is that it does not go far enough. The purpose of this paper is to contend that it is potentially misleading to characterize the relationship between child and adult models solely as one in which the child model successively comes to approximate the adult model as the child grows older. A more productive view is that, while any child model must allow for the developmental fact that the child does become an adult, the child's cognitive response is both quantitatively and qualitatively different from that of adults.

The Role of Developmental Stages

A child model of cognitive response to advertising must of necessity be dynamic rather than static. Any child model must allow for the process of developmental change. This is usually handled through the construct of developmental stages. The use of the stage construct is illustrated in a study by Rossiter and Robertson (1975) which compared sources of variance in children's ability to comprehend the structure and purpose of television commercials. It was shown that children between the first and fifth grade differ reliably in their ability to perform cognitive operations in understanding commercials. The operations were those of classification, discrimination between commercials and programming; relational thinking, perceiving the relationship between commercial messages and an external sponsor; drawing implications, detecting informative intent and persuasive intent; and transformations, dealing with symbolic claims about products. Of the four operations, transformations were the most difficult. But the significance of this research for illustrating the stage construct is that developmental level, as indexed by age and grade, was shown to account for 40% of the total variance--or three-fourths of the accounted for variance--in children's comprehension. This versus 9% for social environment factors and 4% for viewing time differences within grade.

Age itself, however, is not a stage construct, though it can be a useful surrogate variable for research. The passage of time has no scientific status in and of itself. The stage component of a child model must identify the way in which information processing changes over time, both in terms of processing capacity and the

abilities of information acquisition and representation. In its most general form, the stage construct calls for a model of the child at any given developmental level. A model whose relation to a comparable model at any other level is well-defined. The best known and most elaborate stage models are, of course, those of Piaget (cf. Piaget and Inhelder, 1969). Most discussions of the stage component of children's consumer information processing models (e.g., Ward, 1974) cite Piaget's work as a point of reference. Nonetheless, it is not completely clear that Piaget's treatment of stages is suited for such models.

It is important to recognize that Piaget has attempted to formulate a general theory of how the child comes to know about the world. The theory deals with operative knowledge, how the child is able to deal with real events. These events may be task skills, operations, rules, concepts, or beliefs. The basic theoretical idea is that knowledge depends on the development of internal cognitive structures. Events are "assimilated" into the cognitive structure and, at the same time, the cognitive structure "accommodates" to features of the real event. These cognitive structures are not, however, what they are commonly taken for. They are not internalized mediational representations of real events or external information (cf. Furth, 1969). Piaget denies that language is the basis of thought and, moreover, that cognitive structures depend on symbolic or imaginal representations. Piaget's cognitive structures are perhaps best thought of as internal actions. When these actions are integrated into logical systems, they are referred to as operations. By stages Piaget simply means an orderly progression of relatively stable cognitive structures (Furth, 1969). A model at any level is couched in terms of a description of the cognitive structures attained and their relation to the structures of previous levels. At no developmental level are structures considered to be internalized representations.

Note that the Piagetian approach, while extremely interesting, denies the mediational representational view of information processing which is so characteristic of the communications and consumer behavior literature. Indeed, Piaget's approach runs against the historical drift of Western thought in general (Furth, 1969). It may be that Piaget's biological approach and the representational approach are not incompatible, but it seems to us that they are definitely intended for different problems. What one cannot do is impose the notion of a Piagetian stage onto a mediational information processing model. It is not possible to "apply" Piaget's thinking to the typical consumer information processing model. Piaget's cognitive structures are not intended to explain the how and what of the child's cognitive response. These responses may be regulated by structures, or internalized operational schemes, but they are not explained by structural changes. Stages are important only for longer-term, biological maturation.

Consider, for example, the case of perceptual boundedness. This refers to the tendency of children to focus on immediately perceived stimulus features rather than on more delayed or abstracted features. All perceptual boundedness does is to describe a limitation characteristic of a given level of development. It does not explain the

nature of the child's cognitive response.

#### Non-Piagetian Differences Between Child and Adult Models

Recently Calder, Robertson, and Rossiter (1975) have outlined a range of possible non-Piagetian differences between child and adult models. To elaborate further on the differences between child and adult models, as we see them, we will focus here on the child's comprehension, retention, and evaluation of advertising information.

Although our knowledge of what advertising information is salient to children is far from complete, an even more challenging question concerns how such information is represented. What are the basic products of the child's comprehension process? There are at least three types of memory codes, or forms in which information may be stored. Children's research should not follow the pattern of adult research in ignoring all but the linguistic coding of information, although this will certainly be a temptation because of the dominance of questionnaire and rating scale techniques. Considerable emphasis must be placed on enactive (motor) and imagery codes. The presence of enactive codes is perhaps best seen in early development in the form of imitative modeling. Gewirtz and Stingle (1968), for instance, argue that imitation involves learning from a model those behaviors that work--that yield the child what he wants. Verbal imitation also seems to precede linguistic understanding (McNeill, 1970). Children may thus be observed to repeat slogans in a purely enactive, "meaning-less" way (Wells, 1965).

The importance of motor codes probably decreases with development. The child gradually relies more heavily on imagery codes. These differ from symbolic codes in that imagery representation is essentially "stimulus faithful" whereas symbolic representation bears only an arbitrary relationship to external stimulus features. We may speculate that visual imagery plays an important role in children's product related behavior. Take, for example, a child accompanying a parent to a supermarket or toy department. The child might only have information about a cereal or toy in the form of a visual image obtained from a television commercial, but this image may yield a "match" when the child actually sees the product. This familiarity effect could engender preference indications such as pointing.

Imagery occurs in other forms too. Representation in the echoic mode is possible. As already noted, children often repeat brand names, slogans, and jingles from advertisements. This may reflect largely the echoic storage of information. Musical encoding of jingles and themes provides an interesting case. Nonmusical verbal codes are subject to considerable interference from conversation and require specific abilities. But echoic codes are subject to very little interference and involve storage only of the acoustic pattern. Retrieval of this echoic code may influence purchase preference or at least suggest a preference to parents. Again, such a process may not be tapped with verbal research instruments.

In addressing the role of nonverbal stored information we do not mean to suggest that enactive or image codes exclude or necessarily replace linguistic codes. For adults, though imagery may be quite important, much information is probably stored in parallel fashion. Our hypothesis is that children differ from adults in the extent to which these parallel systems are developed. The child's use of linguistic codes is greatly hampered by psycholinguistic factors.

While memory codes are the most important issue for comprehension, the organization of coded information in memory is the central issue for retention and evaluation. Evidence indicates that the dimensions the child uses to organize information stored in memory changes with development. Rossi and Wittrock (1971), for example, investigated changes in children ranging from 2 to 5 years in mental age. Younger children were more likely to use phonetic relationships (rhyming) in memory organization whereas usage of conceptual categories increased with age. We might also hypothesize that the use of simple, linear belief structures predominate with younger children, the use of hierarchical structures perhaps increasing with age. This is especially plausible considering the memory codes employed by children. In any case, the child's capacity to retain information, especially from the media, no doubt depends greatly on his ability to organize the information.

With adult models, evaluation is also thought to depend on how product information is organized. Many evaluation models for adults are best thought of in terms of information integration (Calder, 1975). The notion that evaluation entails some combinatorial process, however, becomes quite tenuous with children. Evaluation for children may be contingent on isolated bits of information which the child happens to be focusing on. There may be no attempt to integrate the stored information. This would explain why children's preferences are probably more unstable than adults! It would lead one to expect recency effects for the presentation of product information too. In short, the summative approach to modeling evaluation cannot be applied uncritically to children.

#### Conclusion

In conclusion, we have alluded to a number of possible differences between child and adult cognitive response to advertising. While all differences may be related to developmental processes, they are not tied, at least in any well-defined way, to the Piagetian construct of stages. They are defined in terms of the child's mediational representation, organization, and evaluation of advertising information.

In addition to suggesting areas for children's research, differences between child and adult models may serve to draw attention to problems which have not yet been recognized in the area of adult consumer information processing. The child and adult areas are complementary. Research on children must be carried out within a theoretical framework emphasizing all aspects of the relationship between child and adult models.

#### References

- Calder, B. "The Cognitive Foundations of Attitudes: Some Implications for Multi-Attribute Models" in M. Schlinger, ed., *Advances in Consumer Research*, Vol. 2 (University of Illinois at Chicago Circle, 1975, 261-268).
- Calder, B., Robertson, T. and Rossiter, J. "Children's Consumer Information Processing," *Communication Research*, Vol. 2 (July, 1975), 307-316.
- Furth, H. *Piaget and Knowledge* (Englewood Cliffs, N.J.: Prentice-Hall, 1969).
- Gewirtz, J. and Stingle, K. "The Learning of Generalized Imitation as the Basis for Identification," *Psychological Review*, Vol. 75 (1968), 374-379.



McNeill, D. The Acquisition of Language (New York: Harper & Row, 1970).

Piaget, J. and Inhelder, B. The Psychology of the Child (New York: Basic Books, 1969).

Rossi, S. and Wittrock, M. "Developmental Shifts in Verbal Recall between Mental Ages Two and Five," Child Development, Vol. 42 (1971), 333-338.

Rossiter, J. and Robertson, T. "Canonical Analysis of Developmental, Social, and Experiential Factors in Children's Comprehension of Television Advertising," Journal of Genetic Psychology, (1975, forthcoming).

Wells, W. "Communicating with Children," Journal of Advertising Research, Vol. 5 (1965), 2-14.

AUTHOR INDEX

Aaby, N. E. ....	58	Hulbert, J. M. ....	405
Abdel-Ghany, M. ....	229	Hunt, H. K. ....	259
Ahtola, O. T. ....	481	Hutt, M. D. ....	36
Albuam, G. ....	88		
Arndt, J. ....	213	Jacoby, J. ....	1, 134, 306, 315
Atkin, D. K. ....	513	Johnson, R. M. ....	353
Axelrod, M. D. ....	437		
		Kakkar, P. ....	370
Barnaby, D. J. ....	246	King, C. W. ....	46
Barnes, J. G. ....	73	Kinnear, T. ....	449
Baron, P. ....	359	Kline, G. ....	290
Barngarten, S. A. ....	457	Kyner, D. B. ....	134
Becherer, R. C. ....	149		
Bernhardt, K. ....	449	Lancaster, K. J. ....	348
Best, R. ....	88	Landon, E. L. ....	263
Bettman, J. R. ....	315	Leavitt, C. ....	252
Beuf, A. H. ....	528	Lee, L. C. ....	506
Beutler, I. F. ....	155, 161	Lipset, S. M. ....	17
Bivens, G. E. ....	229	Lovelock, C. H. ....	407
Blackwell, R. ....	102	Lutz, R. J. ....	370, 469
Bloom, P. N. ....	208		
Bogart, L. ....	12	Markin, R. J. ....	222
Bourgeois, J. C. ....	73	Martin, C. R. ....	453, 463
Burns, A. C. ....	199	McEwen, W. J. ....	341
		Merothra, S. ....	504
Calder, B. ....	536	Miller, K. ....	110
Capon, N. ....	405	Miller, P. ....	290
Chestnut, R. W. ....	134, 306	Mittelstaedt, R. A. ....	30, 58
Cosmas, S. ....	501	Mizerski, R. W. ....	176
		Morris, E. W. ....	161
Day, R. L. ....	263	Morrison, A. ....	290
Denney, W. M. ....	269		
DeVere, S. P. ....	30	Nakanishi, M. ....	24
Dholakia, R. R. ....	387	Narayana, C. L. ....	222
Douglas, S. P. ....	191		
Dover, P. ....	168	O'Shaughnessy, J. ....	364
Dupont, T. D. ....	431	Olshavsky, R. ....	379
		Olson, J. ....	168
Engel, J. F. ....	98		
Eskin, G. ....	359	Painter, J. J. ....	40
Etzel, M. J. ....	485	Park, C. W. ....	184
		Payne, J. W. ....	321
Feldman, S. ....	508	Payne, M. ....	434
Fishbein, M. ....	491	Pekelman, D. ....	81
Fisher, W. ....	306	Percy, L. ....	114, 119, 143
Fitzhenry, N. ....	477	Pitts, R. E. ....	398
Fowles, B. ....	520	Pras, B. ....	92
Glassman, M. ....	477	Rao, T. R. ....	457
Golden, L. ....	63	Ray, M. ....	278
Golob, T. F. ....	416	Recker, W. W. ....	416
Granzin, K. L. ....	40, 68	Reingen, P. H. ....	130
Greeno, D. ....	328	Reizenstein, R. C. ....	246
Grikscheit, G. ....	68	Ring, L. J. ....	46
Grossbart, S. L. ....	30	Ring, L. W. ....	457
		Robertson, L. ....	287
Haines, G. ....	328	Robertson, T. S. ....	508, 536
Harrell, G. D. ....	36	Rosenberg, L. J. ....	261
Hawes, D. K. ....	102	Rossiter, J. R. ....	523, 536
Hawkins, D. I. ....	88	Russo, J. E. ....	273
Hempel, D. J. ....	261, 341	Ryan, M. J. ....	149, 485
Hirsch, P. M. ....	499		
Holbrook, M. ....	364		
Huber, J. ....	138		

AUTHOR INDEX (continued)

Schaninger, C. M.....	184
Schlinger, M. J.....	302
Soctt, C. A.....	387
Sen, Subrata.....	81
Sheth, J. N.....	382, 425
Sommers, M. S.....	328
Sternthal, B.....	387
Summers, J. O.....	92
Szybillo, G. J.....	447
Talarzyk, W. W.....	102
Templeton, J.....	442
Tigert, D. J.....	46
Towle, J. ....	463
Vincent, M. M.....	125
Wackman, D. ....	531
Ward, S.....	278, 531
Wax, S. B.....	276
Weigl, K. C.....	306
Wells, W. D.....	498
Wilkie, W. L.....	334
Williams, F. ....	238
Wilson, D. T.....	394
Wilson, R. D. ....	53
Winter, M.....	161
Woodside, A. G.....	398
Wortzel, L.....	295
Yalch, R. F.....	166
Zikmund, W. G.....	125