

---

# Advances In Consumer Research

---

Volume 2

Mary Jane Schlinger, Editor

---

---

# Advances In Consumer Research

---

Volume 2 1975

Mary Jane Schlinger, Editor  
University of Illinois at Chicago Circle

---

Schlinger, Mary Jane (Editor)  
Advances in Consumer Research, Volume 2

Copyright 1975 by the Association for Consumer Research  
All rights reserved

Library of Congress Catalog Card Number: 74-76437  
ISBN 0-91552-01-9

PREFACE

Advances in Consumer Research, Volume 2 reports the proceedings of the Fifth Annual Conference of the Association for Consumer Research, held November 7-10, at the O'Hare Inn near Chicago.

The Association for Consumer Research is a nonprofit organization established to: (1) provide a forum for the exchange of ideas among those interested in consumer behavior; (2) stimulate research focusing on a better understanding of consumer behavior; and (3) disseminate contributions to the understanding of consumer behavior through seminars, conferences and publications.

As the title implies, the annual Advances in Consumer Research publication provides readers with up-to-date papers from highly qualified researchers in various academic disciplines, government, and business. The many topics covered in this volume reflect the growing scope and diversity of the consumer research field, as well as continuing progress in the methodologies and concepts that are applied to investigations of consumer decision-making.

The 1974 conference was planned and organized by the Association's program committee and officers whose names are listed below.

March, 1975

University of Illinois  
at Chicago Circle



Mary Jane Schlinger  
Editor

1974 Program Committee

1974 Executive Committee

CO-CHAIRPERSONS: Joseph T. Plummer  
Leo Burnett U.S.A.

Mary Jane Schlinger  
University of Illinois  
at Chicago Circle

ARRANGEMENTS: Robert Cecil  
Rob Klugman  
Leo Burnett U.S.A.

MEMBERS: George Brousseau  
National Science Foundation

James Clokey  
Jos. Schlitz Brewing Co.

Joel Cohen  
University of Florida

Harry Davis  
University of Chicago

John Fiedler  
Leo Burnett U.S.A.

Mary Gardiner Jones  
University of Illinois

Arnold Mitchell  
Stanford Research Institute

Robert W. Pratt, Jr.  
General Electric Co.

Scott Ward  
Harvard University

PRESIDENT: William D. Wells  
Needham, Harper & Steers

PRESIDENT-ELECT: Jacob Jacoby  
Purdue University

PAST PRESIDENT: Robert W. Pratt, Jr.  
General Electric Co.

EXECUTIVE SECRETARY: James Taylor  
University of Michigan

TREASURER: Philip G. Kuehl  
University of Maryland

NEWSLETTER EDITOR: Jerry Olson  
Pennsylvania State University

MEMBERSHIP: Robert B. Woodruff  
University of Tennessee

PUBLICATIONS: Stanley D. Shores  
Procter & Gamble

DIRECTORS: M. Venkatesan  
University of Iowa

Harold Kassarian  
University of California at L.A.

Gerald Zaltman  
Northwestern University

What is the answer?

In that case, what is the question?

-Gertrude Stein

## TABLE OF CONTENTS

### STATE OF THE ART: CRITICAL REVIEWS AND EMPIRICAL PERSPECTIVES

Perceived Risk and Consumer Behavior: A Critical Review Ivan Ross	1
Cognitive Dissonance and Consumer Behavior: A Review of the Evidence William H. Cummings and M. Venkatesan	21
A Modern Sociological Approach to the Stratification of Material Life Styles Marcus Felson	33
Life Cycle Revisited: Applications In Consumer Research Subhash C. Jain	39
The Status of Consumer Behavior: Some Empirical Perspectives Gary T. Ford, Philip G. Kuehl, and Robert F. Dyer	51

### CONSUMER DECISION-MAKING PROCESSES

Observing Stages in Consumer Decision Processes Johan Arndt and Edgar Crane	63
Family Role Structure and Housing Decisions Donald J. Hempel	71
Effects of Prior Decision-Making, Demographics, and Psycho- graphics on Marital Roles for Purchasing Durables Arch G. Woodside	81
Intergenerational Patterns of Consumer Behavior Brent C. Miller	93
Methodological Considerations in Family Decision-Making Studies Marilyn M. Dunsing and Jeanne L. Hafstrom	103
Some Unanswered Questions on Family Decision-Making Robert Ferber	113
Behavioral Decision Making: A Comparison of Three Models J. Paul Peter and Lawrence X. Tarpey, Sr.	119
The Use of Seals of Approval in Consumer Decision-Making as a Function of Cognitive Needs and Style Thomas L. Parkinson	133
The Information Seekers--A Cross-Cultural Consumer Elite Jack L. Engledow, Hans B. Thorelli, and Helmut Becker	141

## SHOPPING BEHAVIOR

- A Generalized Utility Model of Shopping Behavior  
Jimmy E. Hilliard, Ronald L. Vaughn, and Fred D. Reynolds 157
- Conceptual and Operational Problems with Market Share Models  
of Consumer Spatial Behavior  
David L. Huff and Richard R. Batsell 165
- Retail Market Area Shape and Structure: Problems and  
Prospects  
J. Barry Mason 173
- A Multivariate Analysis of Prepurchase Deliberation and  
External Search Behavior  
Stephen J. Miller and William G. Zikmund 187
- Credit Riskness of Low-Income Consumers  
Donald E. Sexton, Jr. 197
- Scarcity and Hoarding: Economic and Social Explanations  
and Marketing Implications  
Ronald Stiff, Keith Johnson, and Khairy Ahmed Tourk 203

## LIFE STYLE ANALYSIS

- An Analysis of Moviegoers by Life Style Segment  
Glen Homan, Robert Cecil, and William Wells 217
- Brand Choice and Communication Behavior: Consumer Response  
to the Marketing of Religion  
Donald J. Hempel and William J. McEwen 231

## ATTITUDE THEORY AND ATTITUDE RESEARCH

- The Cognitive Foundations of Attitudes: Some Implications  
for Multi-Attribute Models  
Bobby J. Calder 241
- The Process of Attitude Acquisition: The Value of a  
Developmental Approach to Consumer Attitude Research  
Jerry C. Olson and Andrew A. Mitchell 249
- The Extended Fishbein Model: Additional Insights...  
and Problems  
Micheal J. Ryan and E. H. Bonfield 265
- Product Class Experience, Dimensionality and Reliability:  
Their Relationship in a Nonmetric Scaling Study  
Adrian B. Ryans and Terry Deutscher 285
- The Moderating Effect of Product Knowledge on Multi-Attribute  
Attitude Model Predictions  
Richard L. Oliver 295

ATTITUDE THEORY AND ATTITUDE RESEARCH (cont'd.)

- Perceived Instrumentality and Value Importance of Newspaper Information  
Flemming Hansen 307

COMPARATIVE ANALYSES OF ATTITUDE MODELS

- Ideal Point Versus Attribute Models of Brand Preference: A Comparison of Predictive Validity  
Bernard Dubois 321
- Similarities and Differences of Generalized Brand Attitudes, Behavioral Intentions, and Reported Behavior  
Arch G. Woodside, James D. Clokey, and Joan M. Combes 335
- Explaining Behavior with Weighted, Unweighted and Standardized Attitude Scores  
Sherren Waung 345
- A Multimethod Approach to Validating Multi-Attribute Attitude Models  
James R. Bettman, Noel Capon, and Richard J. Lutz 357
- Multi-Attribute Attitude Models: A Comparative Analysis  
Morris B. Holbrook and James M. Hulbert 375
- Prediction of Attitudes: A Comparative Study of the Rosenberg, Fishbein and Sheth Models  
Secil Tunçalp and Jagdish N. Sheth 389
- Predictive Validation and Cross-Validation of the Fishbein, Rosenberg, and Sheth Models of Attitudes  
P. S. Raju, Rabi S. Bhagat, and Jagdish N. Sheth 405

SITUATIONAL INFLUENCE IN BUYER BEHAVIOR

- The Objective Situation as a Determinant of Consumer Behavior  
Russell W. Belk 427
- The Psychological Situation As a Determinant of Consumer Behavior  
Richard J. Lutz and Pradeep Kakkar 439
- A Situational Multi-Attribute Attitude Model  
Kenneth E. Miller 455

ATTRIBUTION THEORY

- Attribution Theory in Marketing Research: Problems and Prospects  
R. E. Burnkrant 465
- A Test of the Relationship Between Trait and Causal Attribution  
Richard W. Mizerski 471



## MEASUREMENT AND ANALYSIS

Composite Population Descriptors: The Socio-Economic/Life Cycle Grid Richard B. Ellis	481
Awareness as an Indicator of New Product Performance David W. Olson	495
Effects of Polarity of Semantic Differential Scales in Consumer Research John Dickson and Gerald Albaum	507
Bootstrapping of Data and Decisions Joel Huber	515
Service Analysis: A Bank Marketing Example Using Perceptual Mapping William Gillette and Richard H. Evans	525
Application of Discriminant Analysis in Formulating Promotional Strategy for Bank Credit Cards James B. Wiley and Lawrence M. Richard	535
Development of a Scale for Innovativeness Clark Leavitt and John Walton	545
An Empirical Test of a Scale for Innovativeness C. Samuel Craig and James L. Ginter	555
An Instrument to Measure Consumer Self-Actualization George Brooker	563

## ADVERTISING RESEARCH

The Advertising Pretest as Part of a Multimeasure, Multi-method, Multisituation Validation and Application Research System Michael L. Ray	577
Recent Experiences in Copy Research Leland E. Ott	589
Copytesting--Communication vs. Persuasion Fran Kahn and Larry Light	595
Theory as a Bridge between Description and Evaluation of Persuasion Clark Leavitt	607
Preference Types Among Viewers of Toothpaste Commercials Dwight A. Williams	615

ADVERTISING RESEARCH (cont'd.)

- A Corrective Advertising Study  
Harold H. Kassarian, Cynthia J. Carlson, and  
Paula E. Rosin 631

- The Relationships among Advertising, Economic Development  
and Consumer Welfare: An Examination of Recent World-  
Wide Evidence  
Anthony F. McGann and Nils-Erik Aaby 643

EXPERIMENTAL RESEARCH AND THEORIES OF CONSUMER CHOICE

- Introduction: Experimental Research and Theories of  
Consumer Choice  
Penny Baron and Gerald Eskin 655

- Experimental Studies of Consumer Demand Behavior:  
Towards a Technology of Making the Slutsky-Hicks Theory  
Technologically Applicable to Individual Behavior  
Raymond C. Battalio and John H. Kagel 657

- Theories, Facts and the Analysis of Consumer Choice  
Raymond Dacey 671

- Some Conceptual Issues in the Testing of the Economic  
Theory of Consumer Demand  
Penny Baron, Gerald Eskin, and Cliff Lloyd 679

- An Experimental Study of the Relationship Between  
Responses to Price Changes and the Price Level for Shoes  
Nonyelu G. Nwokoye 693

VALIDITY PROCEDURES IN CONSUMER RESEARCH

- Incorporating Tests of Data and Model Validity in Commercial  
and Academic Research  
Edgar A. Pessemier 705

- Detecting Demand Characteristics in Laboratory Experiments  
in Consumer Research: The Case of Repetition-Affect Research  
Alan G. Sawyer 713

- Validity Procedures at the Survey Research Center  
F. Thomas Juster 725

- Validity and Goodness of Fit in Data Analysis  
Donald R. Lehmann 741

- Validity  
Larry Light and Fran Kahn 751

## VALIDITY PROCEDURES IN CONSUMER RESEARCH (cont'd.)

- Discussion: Validity Procedures in Consumer Behavior  
Research  
Albert C. Rohloff 755
- Validity Procedures in Consumer Research: A Perspective  
Albert V. Bruno 757

## MARKETING POLITICAL CANDIDATES

- Overview of Political Candidate Marketing  
Philip Kotler 761
- Images and Voters' Decision-Making Processes  
Dan Nimmo 771
- Attribution Theory and Voter Choice  
Richard F. Yalch 783
- Applications of Marketing Concepts to Candidate Marketing  
Avraham Shama 793

## LEISURE TIME RESEARCH

- Relationships between Consumers' Shopping and Leisure  
Activities and their Attitudes toward the Energy Crisis:  
A Cross Sectional Study  
Glenn S. Omura and W. Wayne Talarzyk 803
- Consumer Satisfaction from Leisure Time Pursuits  
Douglass K. Hawes, W. Wayne Talarzyk, and  
Roger D. Blackwell 817
- Markets for Leisure Time  
Justin L. Voss and Roger D. Blackwell 837
- Time as an Indicator of Social Change and the Quality of Life  
John P. Robinson 847

## PUBLIC POLICY AND PLANNING

- Researching and Modeling Consumer Choice Behavior in Urban  
Transportation  
Christopher H. Lovelock 851
- The Effects of Energy Availability and Costs on Consumer  
Attitudes and Behavior  
Harry W. O'Neill 863
- Perceptions of Children's Television Advertising: An Empirical  
Investigation of the Beliefs and Attitudes of Consumer, Industry,  
and Government Respondents  
James D. Culley 879

PUBLIC POLICY AND PLANNING (cont'd.)

Current Costs of Families' Payments Compared to Estimates of Health Plans Flora Williams	891
Television Advertising: A Study on its Applicability for Promoting the Disadvantaged to Employers James E. Haefner	905
Social Science Inputs to Public Policy Formulation: Massachusetts and the Unit Pricing Regulations Saul Barry Wax	915
Consumer Research of the Office of Energy Conservation and Environment, Federal Energy Administration Jeffrey S. Milstein	925
AUTHOR INDEX	929



PERCEIVED RISK AND CONSUMER BEHAVIOR:  
A CRITICAL REVIEW

Ivan Ross  
University of Minnesota

The empirical research relating perceived risk to consumer behavior is summarized. The literature reveals that perceived risk has been studied in relationship to information acquisition and processing constructs such as word-of-mouth behavior and opinion-leadership, as well as to overt consumer behaviors such as new product adoption, store/brand loyalty, and modes of shopping. Recent research has been concentrated on the study of relationships between specific kinds or components of perceived risk or risk consequences and the specific relievers or reducers of these components. The reviewer offers a critique of research on perceived risk and suggests direction for future research.

Introduction

When Bauer (1960) first proposed that consumer behavior could be viewed as an instance of risk taking, he modestly hoped that the "fad" he was probably introducing would "at least survive through infancy" (p. 23). After fourteen years there is evidence that the infant is fast becoming adult. Indeed, as the list of references suggests, recent years have shown a dramatic increase in publication frequency of empirical research in this area, and current models or theories of consumer behavior broadly incorporate the perceived risk construct. Engel, Kollat and Blackwell (1973) position perceived risk specifically in the "external search and alternative evaluation" stage of decision-making (pp. 376-380) and generally attribute to it great importance: "Decision making (processes) ... occur in order to reduce perceived risk to tolerable levels (p. 59)." Howard and Sheth (1969) conceptually deal with the construct under their term, "stimulus ambiguity", viewed as a "perceptual construct" in their theory of consumer behavior (p. 30).

The reviewer has not found the organization of the empirical literature on perceived risk for the purposes of this paper an easy task nor one which is likely to be optimally satisfying to some (hopefully, not most) readers. In the first place, perceived risk has been studied in relation to a very large number of other consumer behavior variables--too large a number to review in detail within the space limitations imposed. And secondly, the manner in which the construct has been operationally and even conceptually defined has varied so much across the studies, that efforts at synthesis are hampered by questions of "are these two studies really talking about the same thing?" More often than not, the answer is, no.

After a discussion of the conceptualization of the construct, the reviewer has chosen to organize his summary by discussing the major consumer behavior variables to which perceived risk has been applied. Some "problematic" areas in this research tradition and suggestions for future research are at the end of the review paper.

## Conceptualization of Perceived Risk

Bauer's initial proposition was that, "Consumer behavior involves risk in the sense that any action of a consumer will produce consequences which he cannot anticipate with anything approximating certainty, and some of which at least are likely to be unpleasant" (1960, p. 24). Thus, the two primary structural dimensions were uncertainty and consequences which much, but not all, subsequent research in perceived risk has used in the measurement procedure. Bauer strongly emphasizes that he is concerned only with subjective (perceived) risk and not "real world" (objective) risk.

It should be noted that Bauer clearly views perceived risk as not only related to consumers' pre-decision information acquisition and processing activity but to post-decision processes as well. Hence, he describes dissonance theory as concerned with "... ways in which people reduce perceived risk after decisions are made. People will seek out information that confirms the wisdom of their decisions" (p. 32). It would have been well if some researchers in the perceived risk area had more carefully noted the view of dissonance reduction as risk reduction processes who, as a result of failing to do so, drew equivocal conclusions from their research (e.g. Arndt, 1968a; Cox and Rich, 1964; and Schiffman, 1972). In all these cases perceived risk measures were taken after the purchase had (or had not) occurred, at which time it would be reasonable to assume that risk/dissonance reduction processes had begun, and hence would likely contaminate their response to the risk measure. Indeed, these studies might better have been addressed to post-purchase dissonance/risk reduction activity explicitly since we have no examples in the risk literature of such studies.

Cox (1967a) in his initial elaboration of Bauer's conceptualization states that it is often necessary to infer the presence of perceived risk since "... consumers may be unable or unwilling to specify that a situation confronting them is risky (p. 36) ... (thus) ... we will assume, for operational purposes, that risk is, in some way, perceived by our subjects in those situations in which they act in such a way as to handle (e.g. reduce) risk" (p. 37). The amount of perceived risk is construed to be a function of (1) "The amount that would be lost (i.e., that which is at stake) if the consequences of the act were not favorable, and (2) the individual's subjective feeling of certainty that the consequences will be unfavorable" (p. 37). The amount at stake "... is a function of the importance or magnitude of the goals to be attained, the seriousness of the penalties that might be imposed for nonattainment, and the amount of means committed to achieving the goals" (p. 38). Whereas certainty and consequences determine the amount of perceived risk, "The nature of the risk perceived should be a function of the nature of the buying goals involved" (p. 38). Given this "two factor" view of risk structure it then follows that risk might be reduced to a "tolerable level" by either or both (1) reducing the amount at stake (e.g. reducing that which the person hoped to gain, reducing the penalties for failure, and reducing the means by which the gain was to be made) and (2) increasing the degree of certainty that loss will not occur; that is, becoming more certain that action consequences would be favorable.

While most subsequent research has employed these two dimensions specifically (e.g. Cunningham, 1967a), others have used a variant two-dimensional definition such as uncertainty and importance (e.g. Schiffman, 1972; Arndt, 1968b), and some use just one dimension (e.g. uncertainty only, Arndt, 1968a). In some cases it is difficult to distinguish whether uncertainty or consequences is being measured (e.g. "how risky is the purchase of \_\_\_\_"). Bettman (1973)

specifically conceptualizes risk dimensionality as different from that of Cox (1967a) and Cunningham (1967a) by substituting importance for consequences/dangers. "... the risk inherent in a brand choice situation within a product class will depend upon the degree to which a buyer believes he can construct a reasonable decision rule for making a brand choice, and the importance to him of making a satisfactory choice within that product class" (Bettman, 1973, pp. 184-185). And rather than rating the uncertainty directly, Bettman's procedure is to compute the percentage of brands falling above an acceptable level of quality to the consumer (Bettman, 1973, 1974). He reports research (1973, 1974) which supports his conceptualization as opposed to Cunningham's (1967a) but he bases his arguments largely upon his finding that when uncertainty and importance are measured as he proposes, both these components contribute variance to the overall risk ratings, whereas using Cunningham's (1967a) uncertainty and danger components, by far the most variance is explained by the danger component alone. That is, in the Cunningham procedure, when there is a great deal of danger, certainty doesn't matter and effects of certainty are felt only at low levels of danger. In the Bettman procedure, the effects of certainty are most pronounced at high levels of importance, which is what Bettman argues "should" be the case. Bettman (1972) did find that when uncertainty and danger are measured as defined by Cunningham, the two components are not independent, and that the danger component is clearly more important than the uncertainty component (also see Slovic and Lichtenstein, 1968). The issue raised here is not moot, but is very difficult to address empirically, since there may well be differences of opinion in what the conceptual definition of risk is thus leading to different views of its fundamental dimensional structure. Whether or not the relationship between the two dimensions, uncertainty and consequences (or importance), is additive or multiplicative (most have assumed it multiplicative) was tested by Bettman (1972; 1974) as a combination rule or "cognitive algebra" question, and through both graphical and statistical tests he found support for the additive rather than the multiplicative procedure.

Implicit to the questions raised by Bettman is his distinction between "inherent risk" and "handled risk." "Inherent risk is the latent risk a product class holds for a consumer, the innate degree of conflict the product class arouses in the consumer. Handled risk is the amount of conflict a product class engenders when the buyer chooses a brand from that product class in his usual buying situation. Thus, handled risk includes the effects of information and risk reduction processes as they have acted on inherent risk" (Bettman, 1972, p. 394). Bettman notes that these two different types of risk have been confused in the research literature; Cunningham (1967a) using inherent risk and Cox and Rich (1966) and Spence, Engel and Blackwell (1970) using handled risk. Subsequently, only a few researchers (e.g. Lutz and Reilly, 1973) have explicitly noted the type of risk they are measuring in the sense of Bettman's distinction. That the distinction is important is demonstrated by Bettman's (1972) research which found that of nine products studied, toothpaste and margarine had the highest ratings for relative inherent risk while beer and instant coffee had the highest relative ratings for handled risk.

#### Perceived Risk, Word-of-Mouth, and Opinion Leadership

Word-of-mouth and opinion leadership were the concepts first researched in relationship to perceived risk, perhaps because Bauer had asserted that "... one of the very important functions of opinion leaders is to reduce the perceived risk of the behavior in question" (Bauer, 1960, p. 26). Research on the manner in which physicians adopted new drugs undoubtedly influenced his view (Coleman, Katz, and Menzel, 1957; Coleman, Menzel, and Katz, 1959). This research suggested



that doctors tended to rely on their colleagues, especially the more "respected" ones, early in the diffusion process, rather than on non-professional medical sources. Once the drug had become reasonably well established, personal influence seemed to play a less important role. As the severity of the disease for which the drug was to be used increased so did the propensity for doctors to rely upon professional as compared to commercial sources (Bursk, 1960; Bauer and Wortzel, 1966).

Cunningham (1964, 1966, 1967a, 1967b, and 1967c) measured the uncertainty and danger (consequences) housewives perceived in the fabric softener, dry spaghetti, and headache remedies product categories (uncertainty -- would an untried brand work as well; consequences -- how much danger would she see in trying a brand she had never used before.) In addition, brand purchase behavior, word-of-mouth activity, and various descriptive measures were obtained. Cunningham hypothesized that "... those users of a product who were high in perceived risk would reduce risk through conversation and thus a greater proportion of the high risk perceivers would be classified as 'talkers' (a respondent who discussed the product category within the last six months) than would the low risk perceivers" (1967b, p. 271). The data supported this hypothesis for headache remedies and fabric softeners but not for dry spaghetti. Regarding the direction of flow of word-of-mouth as a function of perceived risk, there were product differences. Those who perceived high risk regarding headache remedies were more likely than low risk perceivers to have initiated their last conversation about the product, but the relationship was reversed for fabric softeners. Also, those higher in perceived risk for fabric softeners and dry spaghetti (but not headache remedies) were more likely to have requested information than those lower in perceived risk and were more likely to claim they had made a recommendation in their last conversation, a finding at variance with the researcher's hypothesis. Nevertheless, Cunningham views this finding as one which "... strongly supports the notion advanced ... that high perceived risk consumers are sought out by others who presumably value their expert opinion" (1967b, pp. 282-283). Thus, in his view of the data, the high risk perceiver is more apt to be an opinion leader. A problem in this research design is that subjects were retrospectively reporting their role in the word-of-mouth process. The question regarding opinion leadership was, "When you bring up the subject of products and brands, do you usually ask someone else for information or do you just suggest helpful information from your own experience?" (1967b, p. 279). Those who are high risk perceivers might report that they suggested information to others (more than low risk perceivers) as part of a risk (dissonance) reduction process. Cunningham recognizes this issue but discounts this interpretation in sustaining the conclusion reported above.

Arndt (1967b, 1967c, 1968b, 1968c) studied word-of-mouth flow within a married student housing complex concerning the adoption process for a new brand of coffee, PERKY, and the data support his hypothesis that opinion leaders would be lower in perceived risk, contrary to Cunningham's research findings above. Arndt believes that the differences between the two studies might be explained by differences in methodology or in products chosen for study. Indeed, there were methodological differences; Arndt measured importance ("How important is it to you that a new brand of coffee you have never tried before is as good as your present brand: not important, fairly important, important, or very important?") rather than consequences. Arndt also found that word-of-mouth had more effect on high- than on low-risk perceivers. That is, those who were high risk perceivers appeared to pay more attention to what they had heard, particularly to unfavorable comments. In general, Arndt views his data as supportive of the conclusion that, "word-of-mouth seemed to flow from the low- to the high-risk perceivers" (1967c, p. 294), primarily because the low-risk

perceivers were more likely to report having given advice about PERKY than high-risk perceivers. In most all other regards, however, the high-risk perceivers seemed to be more "active" than the low-risk perceivers: in initiating pre-purchase conversation, in overhearing comments, and in seeking information.

Various other studies have asked subjects to evaluate the importance of alternative information sources, and personal sources (i.e., word-of-mouth) are invariably rated high in importance, and there is evidence that it is in particularly "high-risk" situations where personal influence is most important, supporting Arndt's view (1967a). For example, Roselius (1971) found that those who perceived high risk for "time", "ego", or "money" loss rated word-of-mouth more helpful (as a "reliever") than did subjects in general. And Perry and Hamm (1969) found that "... the higher the risk involved in a particular purchase decision, the greater the importance of personal influence" (p. 354) in their study of social and economic risks perceived by subjects across 25 product categories. The same conclusion was reached by Lutz and Reilly (1973) in their study of the effects of social and performance risk on consumer information acquisition; word-of-mouth was the most important of the four sources of information available externally to subjects (word-of-mouth, mass media advertising, rating magazine, and sales clerks). However, they did not find that variations in levels of social risk influenced information search behavior as they had hypothesized. Finally, Sheth and Venkatesan (1968) also found that the experimentally created high-risk (regarding the hair-spray product category) group sought personal sources of information significantly more than did the low-risk group.

The research on perceived risk, word-of-mouth, and opinion leadership would seem to support the generalization that word-of-mouth functions as an important (but not necessarily the most important, e.g. Roselius, 1969) risk reliever across most or all types of risks. The nature and direction of word-of-mouth flow, and specifically as this relates to opinion leadership, is less clear. Certainly this is a very complex issue and one not easily amenable to investigation, especially through a self-report mode.

#### Perceived Risk, New Product Adoption and Brand/Store Loyalty

The drug adoption studies previously referred to (e.g. Coleman, et. al., 1957; and Coleman, et. al., 1959) suggest the hypothesis that those high in perceived risk for a product category would be less likely to adopt at all, or to adopt quickly, if at all, a new brand introduced within that category, and vice-versa. The research subsequently would seem to strongly support this hypothesis. Both Arndt (1967b) and Cunningham (1967b) in the studies previously referred to, found evidence to this effect (although not necessarily a clear relationship for all products studied), as did Schiffman (1972) in his study of the adoption of a salt substitute product among elderly consumers. If one accepts Sheth's (1968) assumption that the adoption of a stainless steel razor blade is a low-risk decision, then one might take his results as evidence for the same hypothesis: "... as high as 89 percent of total respondents adopted the stainless steel blades in slightly more than a year's time from the influx of three major brands in the market in 1963" ... and ... "more than 90 percent of total respondents adopted within one year after becoming aware. Similarly, as high as 49 percent adopted the stainless steel blades immediately after they became aware" (pp. 180-181). He continues, "... it can be easily seen that for a low risk innovation also possessing strong relative advantage, the diffusion is faster both in terms of time of adoption and the mental process of adoption" (p. 181).

The hypothetical relationship between perceived risk and brand/store loyalty is closely related to its relationship with new product adoption. Loyalty should be stronger among those perceiving high-risk in the product category and for basically the same reason: "Much brand loyalty is a device for reducing the risks of consumer decisions" (Bauer, 1960, p. 25). Arndt (1967b) found that those high in perceived risk for coffee were more likely to be brand loyal and hence less likely to adopt the new coffee under study. Cunningham (1967c) similarly found supportive evidence for this relationship but it was less strong for dry spaghetti than for fabric softeners or headache remedies, thus suggesting that where risk is generally low for the product category (e.g. dry spaghetti), brand loyalty plays a smaller role as a risk reduction process. Sheth and Venkatesan (1968) studied the development of brand loyalty as a risk-reduction process in repetitive (over time) consumer behavior and found support for their hypothesis that brand loyalty increased over time. It should be noted, however, that the development of brand loyalty (repeat selection of brands) was quicker for the low-risk than for the high-risk groups. The authors conclude that "...perceived risk is a necessary condition only for the development of brand loyalty. The sufficient condition is the existence of well-known market brand(s) on which the consumer can rely." (p.310)

Hisrich, Dornoff, and Kernan (1972) hypothesized that, "If the product is intolerably ambiguous, perhaps the store, which might be far less ambiguous, can serve as a surrogate" (p. 435). They chose draperies, furniture, and carpeting as their "ambiguous" products, but found their data rejected the hypothesis. For all three products, and for both male and female subjects, and at every level of perceived risk, the number of store-loyal buyers was less than the number of non-loyal buyers. They conclude, "At a minimum, this suggests that these buyers did not consider repeat patronage as a viable risk-handling strategy. Indeed, depending on prior results, not shopping at a previously-patronized store might have served as a form of risk reduction" (pp. 438-439).

#### Perceived Risk and Mode of Shopping

Noting that many women do not order any merchandise by phone, Cox and Rich (1964) hypothesized that telephone shopping creates perceived risk of sufficient magnitude to deter many women from shopping by this mode. Although in their measurement procedure, in the reviewer's opinion, there was contamination between the criterion measure (whether or not the item was purchased by phone) and the dependent variable measure (what items could be bought by telephone without worry and which would be worried about), the authors conclude that "...high perceived risk is likely to be a strong deterrent to purchasing an item by telephone" (p. 499). When respondents were asked why they had not shopped, nearly two thirds replied "...that they had not done so because they were apprehensive of not getting what they wanted" (p. 495). Among those who did shop by phone, newspaper advertising was a favored source of information. However, that newspaper advertisements function to identify the merchandise and the store at which it is available may confound the relationship between "saw ad" and "shopped by phone" so that it might not necessarily suggest anything about the uncertainty reducing role that such advertising might play. That is, ipso facto, the person who does shop by phone as contrasted with one who does not is more apt to say she "relies" more on newspaper advertisements. In any case, the author's investigation of those items of merchandise rated higher and lower in perceived risk in shopping for them by phone suggested that, "The more decisions to be made in making a single purchase, the more important the decisions are, and the more uncertain the consumer is about making the decisions without visual inspection, the greater the risk potential of ordering the product by phone" (p. 505).

Using a paper-and-pencil questionnaire, subjects rated the overall risk they perceived in purchasing twenty different products through the mail as opposed to in a store or from a salesman in research conducted by Spence, Engel, and Blackwell (1970). Their hypothesis that people perceive more risk in buying by mail than buying in a store or from a salesman was supported, but their hypothesis that mail-order buyers of hospitalization insurance would perceive significantly less risk in mail-order buying of other products was not supported by the data, nor was there support for the hypothesis that mail-order buyers of hospitalization insurance would perceive significantly less risk in the mail-order purchase of such insurance than non-buyers. The authors recognize the inconsistency between the support of the first hypothesis and the lack of support for the second and third and urge that future research be directed at the question. It may well be that the overall risk measure employed obscured particular types of risk being more prominent in one mode than in the other, particularly as these risk components might have interacted with the particular products studied. One might further question the reliability of the difference scores (between perceived risk in the two buying situations) computed by the researchers and then subjected to an ANOV, as well as the logic in summing these difference scores across the twenty products for each subject to obtain an "average perceived risk difference" score.

An issue not addressed in either of these studies is the question of the use of different shopping modes in gathering information rather than the more narrow question of actual purchase by that mode. It would be interesting to know whether or not high risk perceivers for buying by telephone are more or less inclined to gather information by phone as compared to lower-risk perceivers. It would also be enlightening were research conducted which simultaneously evaluated the way in which these alternative shopping modes would be used by consumers in coming to a decision rather than addressing the question one mode at a time. For example, one might create an environment where consumers could choose among these modes, each with fixed costs (e.g. postage, gasoline, time, etc.), in deciding on a purchase.

#### Perceived Risk and the Relievers of Perceived Risk

Cox (1967a) very early in the perceived risk research literature noted, at least for the two consumers in his study, that reducing uncertainty was far more common than reducing unfavorable consequences as a risk reducing strategy. Sheth and Venkatesan (1968) state that, "Generally, the consumer cannot change the consequences of using a brand. He can, however, change his uncertainty about these consequences ..." (p. 307). Although there is no research directly related to this question, subsequent researchers have restricted themselves to "uncertainty reducing" strategies. However, to the extent that the certainty and the consequences dimensions of perceived risk are not independent (see Bettman, 1972), we might construe that much of the empirical evidence is functionally addressed to both dimensions simultaneously.

The reviewer has already noted one example of an instance when different levels of overall risk seem to evoke different "relievers" or risk-reducers than when risk is low; namely, the apparent important role of word-of-mouth or personal sources in general. Arndt (1967b) found that the content of perceived risk was different for high versus low risk perceivers for coffee. Those who perceived low risk denied any problem except "inconvenience" but those high in perceived risk saw "waste of money" and "husband's reaction" as risk factors. And in his study of the persuasibility of purchasing agents and chemists, Levitt (1967) found that although a high credibility company source

acted to "reduce risk" in both a "high risk" (adoption of product) and a "low risk" (refer the product to someone else for serious consideration) situation, credibility was more important in the high risk situation. Thus, there is some evidence that the level of overall perceived risk might evoke different prepotencies of risk components, hence different relievers for that risk. It should be noted that research is not consistent on this point. For example, Zikmund and Scott (1973) using a multiple discriminant analysis to distinguish differences in information search activity between high versus low risk perceivers regarding lawn furniture, color TV, and stationery, did not find a statistically significant difference in these information preferences. (However, a canonical analysis performed within product categories did reveal differences.)

But most research has sought to specifically relate types of risk to types of relievers. A starting place seems to have been the difference in evaluation of product information as a function of the "performance" versus "psychosocial" goals of the consumer (e.g. Wilding and Bauer, 1968; Ross, 1972). Subsequently, most recent research in perceived risk has been focused on the relationship between risk and reliever relationships (Roselius, 1971; Perry and Hamm, 1969; Lutz and Reilly, 1973; Zikmund and Scott, 1973; Jacoby and Kaplan, 1972; Kaplan, Szabillo and Jacoby, 1974; Newton, 1967; and McMillan, 1972).

Although they did not relate their risk components to types of relievers, Jacoby and Kaplan (1972) did address the fundamental structure of these components. They identified five types of risk: (1) financial, (2) performance, (3) physical, (4) psychological, and (5) social. Considering the way in which these components grouped themselves as subjects rated these risks with regard to twelve diverse consumer products, it was clear that price seemed to be the metric ordering these products on overall risk perception. Performance risk correlated most highly with overall perceived risk more highly than any other component for eight of the twelve products, and was highly correlated for the others. For that reason, the authors suggest that performance risk could be employed as an approximation of overall perceived risk. However, this generalization may be unwarranted since most of the twelve products "seem" highly "performance" related (performance risk -- "What is the likelihood that there will be something wrong with an unfamiliar brand of \_\_\_ or that it will not work properly"), and thus high correlations between this component and the overall measure ("On the whole, considering all sorts of factors combined, about how risky would you say it was to buy an unfamiliar brand of \_\_\_") may be variously interpreted. In any event, after performance risk, the next most important risks averaging across the twelve products are (in decreasing order) financial, social, psychological, and last, physical. But in regressing components on the overall perceived risk score, social entered after performance risk. The authors found that a multiple regression equation predicting overall perceived risk from component risk scores accounted for 74 percent of the variance in this criterion, and moreover, they cross-validated the regression weights derived in this study using different subjects two years later (Kaplan, Szabillo, and Jacoby, 1974) and found negligible shrinkage. The cross-validation is to be applauded since it stands out as a singular event of its kind in all the perceived risk research literature.

Roselius (1971) identified another type of loss, time loss, and studied the effect of this type of loss in comparison to three others: hazard loss, ego loss, and money loss, the latter three seemingly comprising the same "set" of losses or risks as those employed in the research by Jacoby and Kaplan (1972). There were eleven risk or loss relievers studied with respect to these losses using a five-point rating from 472 housewives on "how helpful (almost always, usually, rarely, almost never) each reliever would be for reducing the risk posed in the

situation": (1) endorsements, (2) brand loyalty, (3) major brand image, (4) private testing, (5) store image, (6) free sample, (7) money-back guarantee, (8) government testing, (9) shopping, (10) expensive model, and (11) word-of-mouth. We are told only that the questionnaire "presented several generalized risky buying situations" in which "situations were not related to specific products or purchase methods," so there is no way for the reader to know the meaningfulness of the stimuli employed. Of the eleven relievers evaluated, "brand loyalty" and "major brand image" evoked the most consistently favorable response, being ranked first and second, respectively, as relievers for each of the four losses. (The authors constructed a "net favorable percentage" quotient for each reliever which was the number of unfavorable responses subtracted from the number of favorable responses given to it by subjects, the difference divided by the total number of responses, then multiplied by 100). Some relievers were consistently rated "unfavorably" by respondents; e.g. "expensive model" was least helpful in all four kinds of losses and had a negative quotient sign, and "private test", "money back guarantee", and "endorsements" all had negative signs across the four losses. One may well question the meaningfulness of a "negative" quotient sign in that it suggests a reliever was "not helpful" (when it could on the average have been helpful) or "unhelpful" (which was not a response alternative for subjects). Some relievers had "special meaning" in the sense that high loss perceivers were particularly sensitive to them; for example, as previously noted, "word-of-mouth" was a helpful reliever except for "hazard loss" ("Some products are dangerous to our health or safety when they fail"), and "major brand image" seemed to function much the same as "word-of-mouth." "Government testing" was found to be particularly helpful as a reliever for hazard loss. Six of the relievers were labeled "general-purpose" risk relievers by Roselius since there was not a significant difference in responses to them between high perceivers and "other" buyers across the types of loss; these were brand loyalty, private testing, shopping, endorsements, expensive models, and money-back guarantees. Roselius concludes that "... buyers prefer some relievers to others depending upon the kind of loss involved ... (and) ... perhaps a seller should first determine the kind of risk perceived by his customers and then create a mix of risk relievers suited for his combination of buyer type and loss type" (p. 61).

Lutz and Reilly (1973) specifically undertook an investigation of the effects of social and performance perceived risk on information acquisition. Subjects rated which of five types of information they would use in making a purchase decision for each of nine products selected by pretest to represent different levels of social and performance risk combinations. They found that when performance risk was low or moderate, subjects opted for "buy" (take their chances and pick a brand without search for product information), but when products were high in performance risk, "direct observation and experience" was the most preferred route and "buy" the least. Over all levels of performance risk, word-of-mouth was the most important of the four information sources external to the consumer (following "buy" and "direct observation and experience"), but contrary to the hypothesis, variations in social risk level did not have any influence on consumers' information search behavior. In a similar sort of research design, Perry and Hamm (1969) tested the effect of social risk and economic risk on subjects' evaluations of seven alternative sources of information and found that "... the higher the risk involved in a particular purchase decision, the greater the importance of personal influence" (p. 354).

Zikmund and Scott (1973) conducted personal interviews with housewives to evoke specific risk consequences associated with purchasing in eight different product classes. These risk variables were then factor analyzed to identify principal risk dimensions. They also measured the traditional uncertainty and consequences of perceived risk and computed an overall risk measure by

multiplying these two components. Considering color TV, lawn furniture and stationery, the factor analysis showed that all three products had "... dimensions relating to quality or reliability and the reaction of significant others who might judge a purchase" (p. 410). A "new" risk factor, not previously identified, was "future opportunity lost", which is associated with the expectation that an improved or lower cost product may be available at a future time which would be precluded by a current purchase -- both color TV sets and lawn furniture have this risk associated with them. They also found a "shopping frustration" factor associated with lawn furniture and stationery but not with color TV. The authors concluded that the research "... illustrates an important reason for investigating perceived risk in a multidimensional fashion. Consumers evaluate products on the basis of a few principal attributes and each represents a potential source of risk. Further, these attributes vary across product classes. Disaggregating perceived risk into product-specific components in this fashion provides much more information about why a consumer perceives risk than overall measures such as social or performance risk" (p. 411).

### Perceived Risk and Personality

Cunningham (1967a) had suggested that some people have a generalized tendency to perceive either high or low risk across a range of products, but this hypothesis has not been specifically tested subsequently. Cox (1967a) noted a "clarifier"- "simplifier" cognitive style difference in the two subjects he studied intensively; the one tending to seek information to clarify or reduce ambiguity ("clarifier") and the other reducing ambiguity by keeping out disturbing cognitive elements ("simplifier").

The perceptual/cognitive style construct, "category width", or "broad" versus "narrow" categorizers, has been specifically studied in relation to perceived risk. Pettigrew (1956) observed that "broad categorizers seem to have a tolerance for type I errors: they risk negative instance in an effort to include maximum positive instances. By contrast, narrow categorizers are willing to make type II errors. They include many positive instance by restricting their category ranges in order to minimize the number of negative instances (p. 532). Popielarz (1967) reasoned that broad categorizers would express greater willingness to buy new products than narrow categorizers and that people with broad category ranges would be more likely to perceive smaller qualitative differences between products of a given product class than narrow categorizers. For each of six products, subjects indicated their willingness to buy each of four qualitatively different brands; the brands differed in newness of products themselves and in the buyers' familiarity with the brand name of the manufacturer (two levels of newness and two of familiarity). They then rated the extent to which they saw brands as qualitatively different. Category width scores correlated with willingness to buy in the hypothesized direction, but only for male subjects was the prediction supported that broad categorizers would perceive smaller qualitative differences among products; for female subjects, the relationship was reversed.

Schiffman (1972) also found that the broad categorizer was more apt to have adopted a new product (salt substitute) than the narrow categorizer, but the measure he employed to measure category width (he refers to the construct as "error tolerance") seems to the reviewer to be criterion contaminated (see "The criterion and construct definition problem" section of this review). And although not specifically related to the perceived risk concept, Barach (1969) found that broad categorizers were more persuaded (i.e. "switched" more to the advertised brand in a Schwerin test) than were narrow categorizers.

In sum, there is evidence that the perceptual or cognitive construct, category width, is related to willingness to adopt new products; broad categorizers being more willing than narrow categorizers to adopt. However, the construct has not been specifically related to perceived risk. We do not know that broad categorizers are less prone to perceiving risks as a generalization.

Self-confidence as a personality construct has also been studied in relationship to perceived risk. Although Hisrich, *et. al.* (1972) found a significant inverse relationship between perceived risk and generalized self-confidence, Zikmund and Scott (1973) and Cunningham (1967a) found no relationship between the two. However, Cunningham (1967c) does report that those subjects medium in self-confidence were more likely to be brand loyal (high brand commitment) than those low or high in self-confidence. The effect of self-confidence on perceived risk remains unclear.

#### The Criterion and Construct Definition Problem

Several investigators have recognized that the criterion problem has not been adequately addressed (e.g. Spence, *et. al.*, 1970, p. 369). Unless we can "know" what kinds of behaviors are manifestations of risk then it is always equivocal that we are really (validly) measuring risk. The criterion problem is at the same time a construct definition problem, and vice-versa. Researchers have assumed that there is risk in decision-making simply because, using the instruments they have developed to measure it, they have "measured" it. Naturally, one could make the same observation about research difficulties associated with other similar hypothetical constructs or intervening variables such as "personality", "self-confidence", "dissonance", "attitudes", and so on. But the reviewer would suggest that the criterion/construct definition problem in perceived risk research appears to be more troublesome and less adequately dealt with than in programmatic research in these other areas. One can find numerous examples among the studies reviewed herein where the "criterion contamination" problem is severe, rendering "findings" of many of these studies equivocal, at best.

For example, Arndt (1968a) measured (and thus defines) perceived risk by asking housewives, "How sure would you be of picking the best brand of (product class): very sure, quite sure, sure, not too sure, or unsure?" (p. 3). A criterion variable is "innovativeness" which he measures by asking when the consumer made her first purchase of the product (ranging from not at all to three or more years ago). He not surprisingly finds a strong negative correlation between perceived risk and innovativeness. Now if a consumer has never bought any brand in the product category (and in his study, 40 percent had never bought soft margarine and 74 percent had never bought electric toothbrushes or electric dishwashers) one should not be surprised to find that she is rather "unsure" that she would be able to "pick the best brand" in that product category.

Cox and Rich (1964) have a similar contamination issue wherein they measure perceived risk in telephone shopping by asking respondents to sort 20 cards, each bearing the name of a particular item of merchandise, into two piles: (1) "Items you feel could be bought by telephone without too much worry over getting just what you want"; and (2) "Items which you would worry about if ordered by phone" (p. 499). All respondents in this portion of the study had ordered by phone at least once during the year prior to the survey. The 10 items about which the respondents had the most "worry" were designated



"high perceived risk" items and the other 10, the "low perceived risk" items. The authors then determined the relationship between the perceived risk ratings merchandise items received and the frequency with which an item was mentioned as being purchased by the telephone in their last phone orders. They find that "... knowing that an item is high in perceived risk allows us to predict that in 90 percent of the cases the item will be a medium or low frequency of telephone purchase item. Knowing the item is low in perceived risk allows us to predict that in 85 percent of the cases the item will be medium or high in frequency of phone purchase." (p. 499) Although the authors caution that "... it is not possible to demonstrate the direction of causality ..." that, "... it would seem reasonable to conclude that high perceived risk is likely to be a strong deterrent to purchasing an item by telephone" (p. 499). One might argue that this observed relationship demonstrates that people say they are apt to be less worried about doing something they have in fact done before than they are to worry about doing something they have not done before. The construct of perceived risk has become obscured by defining (measuring) it in a situation-specific context (ordering by telephone) which is the same context used to "validate" and interpret the effect of the construct. The logic here (or the lack of it) is akin to the assertion that "not having a telephone is likely to be a strong deterrent to purchasing an item by telephone." Although probably "correct", we have not learned much about why some people shop by telephone more than do others.

Contamination problems also seem present in Schiffman's (1972) study of new product trial (a salt substitute) among 100 elderly (average age, 74) consumers. All households had received a coupon worth 30 cents on the purchase of the salt substitute regularly priced at 59 cents. After two weeks, 17 percent had redeemed the (household coded) coupon, soon after which time interviews were held with the female member of all households. "Taste risk" and "health risk" were measured as well as the importance of each of these two types of risk as follows (p. 107): "Would you say you are quite certain, somewhat certain, or not certain that a new brand of salt substitute would taste as good as regular salt (Taste Risk)? And, "Would you say it is not important, of some importance, or quite important for you (or your husband) to get a salty flavor into your food (Taste Risk Importance)? For Health Risk the question was, "Would you say you are quite certain, somewhat certain, or not certain that a new brand of salt substitute would be better for one's health than regular salt?", and the Importance of Health Risk was evoked by, "Would you say there is no danger, some danger, or much danger in using a new brand of salt substitute in place of regular salt?" Since these questions were asked after adoption had occurred, and assuming that taste and health considerations were indeed primary (as author's pilot study had indicated), then we have a backward contamination built into the design in that the retrospective recollection of "risks" among those who had already adopted the salt substitute would be expectedly "lower" than the risk perception of those who had rejected the product, presumably for the very risk "reasons" contained in the risk measure. Further contamination occurs between Schiffman's measure of respondents history of new product adoption (how many of ten new food products introduced in the 18 months prior to the study had the household tried) and his measure of "perceived error tolerance" (another way of saying broad versus narrow categorizers) which was measured by the questions: "Who is a wiser consumer : (1) a person who tries a new food product which turns out to have a poor taste or (2) a person who does not try a new food product and later learns it has a good taste?", and, "Do you prefer to: (1) try a new food product when it first comes out, or (2) wait and learn how good it is before trying it?" (p. 107). The strong positive correlation between the "inclusion" (broad) style and the criterion measure, the number of new food products adopted, is a clearly contaminated relationship

which only demonstrates that if a person knows they've tried new food products lately they will say they try new foods when they first come out. What this demonstrates about even the existence of a "new risk-handling variable, perceived error tolerance", quite aside from the question of what such a construct could or does have to do with new product trial, is unclear.

#### The "Leap" to Conclusions Unwarranted by the Data

This age-old problem is not escaped in the perceived risk empirical research tradition; notwithstanding the "aiding and abetting" of this propensity by journal editors who demand "marketing management implications" to that which they publish, one has reason to be critical of such conclusions, interesting as they may be, especially since many of us who skim the journals almost always read the "findings" and/or "conclusions" sections if nothing else. We would be led astray by many of the conclusions we would read in this area.

For example, Sheth (1968), despite his caveats that the empirical research on the diffusion of innovations has been lacking in theoretical foundation and rather has proceeded "on the grounds of convenience and ease in implementation" (p. 175), and that in the study of diffusion it is important "to consider the adopter's perception of the magnitude of risk involved in an innovation", concludes after his research on the adoption of stainless steel razor blades among 601 college males that "... it can be easily seen that for a low risk innovation also possessing strong relative advantage, the diffusion is faster both in terms of time of adoption and the mental process of adoption" (p. 181). In fact, Sheth does not measure whether or not or the extent to which risk was perceived by adopters but simply asserts that this would be an example of a "low risk" innovation. Further, we know nothing about subjects' perception of "relative advantage" of this product nor anything about the "mental process" of adoption (or non-adoption). Although Sheth may certainly be correct in his theoretical argument that the adoption process for a "miracle drug" may be quite different than that for a stainless steel razor blade in part because of the differences in "perceived risks" among adopters for these products, we find no data in this study which are relevant to this very important issue. If speed of adoption/diffusion is all we're interested in, then why not simply contrast sales curves at retail for razor blades, new drugs, or whatever? Why bother asking adopters when they adopted if this is not specifically to be related to some measure of risk perception?

Perry and Hamm (1969) specifically measure only two risk components, "social" ("how the purchase decision will affect the opinion other people hold of the individual") and "economic" ("how the purchase will affect the individuals ability to make other purchases") (p. 351), and yet draw conclusions with respect to "high-risk purchase situations" despite the fact that they did not include other components of risk in their conceptualization. They also conclude that "These findings suggest that promotional strategies in a high-risk purchase situation ... should ... emphasize the social benefits of the purchase more than the economic ones" (p. 354) on the grounds that the social risk component contributed more variance to total risk scores than did the economic risk component. But the variance finding alone does not suffice as a basis for arguing that "social benefits" would be more "risk reducing" than "economic" ones, although the finding does suggest an interesting hypothesis which ought to be empirically tested.

## Future Research Needs

### 1. Unobtrusive measurement or at least disguised measures of overt risk-reduction processes:

Since the usual research designs have measured risk perception simultaneously with risk relieving preferences or activities in a paper-and-pencil or other self-report mode, the subjects' "set" to be rational and to give "proper" answers is a likely bias which may, for example, account for the relatively low importance attributed to media advertising as a risk-reduction information source. This research mode also is likely to sensitize subjects to their "perception" of perceived risk and may therefore motivate them to behave as if there was risk when, without the intrusive measure, they might not have done so. Admittedly, there is a danger in separating in time the measurement of risk (less in the "inherent" than in the "handled" sense, probably) from the measurement of risk handling processes, but to the extent that it can be done, the two measurements should not be so "nakedly" exposed to one another. Embedding either or both within other measures or other disguises are recommended. Unobtrusive measures have obvious advantages where the subject of the investigation would seem to be so especially sensitive to these potentially reactive measures. Certainly the subject's preferences for and utilizations of some information (e.g. advertising, salespersons, etc.) could be unobtrusively measured by the ingenious researcher.

### 2. Experimental manipulation of risk:

With the exception of research by Sheth and Venkatesan (1968), there is no research in this area which directly manipulates risk by experimental design (although several studies expose subjects to different risk "sets" in evoking preferences for alternative kinds of information, e.g. Hart, 1974). There can and should be more research experimentally manipulating products/services and purchase/use situations (e.g. use by purchaser only versus use by others, purchase by telephone versus in-store, etc.). Use of such designs would give the researcher more power, especially in addressing theoretical relationships within the risk model (e.g. kind of risk perceived and preference for risk relievers).

### 3. Mathematical formalization of risk structure:

As noted by Nicosia (1969) in his review of the Harvard studies (Cox, 1967c), the perceived risk literature "...reflects the shortcomings of a too direct dialogue between verbal hypothesis and empirical data (and their statistical manipulations) without the benefits of an intervening formalization" (p. 165). There is no true model of the perceived risk construct as it relates specifically to information acquisition, transmission, or handling. Taylor's (1974) efforts may be helpful in this regard as well as some of the work on risk enhancement strategies as stimulated by Berlyne (1965) reflected especially in Copley and Callom (1971).

### 4. Risk enhancement strategies:

Berlyne (1965) has asserted that increasing response conflict can be as important as attempts to reduce conflict, especially in monotonous environments wherein persons may engage in "diversive" exploration (p. 244). Deering and Jacoby (1972) and Copley and Callom (1971) have made provocative initial efforts in studying the conditions under which such risk enhancing

activity may occur, and the question seems worthy of further study. Deering and Jacoby specifically hypothesized that "Given purchase alternatives which encompass a wide range of risk, maximal preference should be manifested for alternatives which are neither extremely high nor low in perceived risk" (p. 406) and found some support for this hypothesis but the relationship was more complex than expected. Copley and Callom studied industrial search behavior and did identify a group which behaved as the "Berlyne curve" would suggest; however, only 8 percent of the subjects in their study behaved in this way. And Venkatesan's (1973) work in "novelty-seeking" may be a conceptually related construct which ought generate empirical research in consumer behavior.

5. The literature reviewed suggests that perceived risk is a function of intra-personal variables (e.g. personality), product differences, and situation differences (e.g. for "self" or for "other" -- see Hart, 1974; and Reingen, 1974). There may well be demographic correlates but the research obscures such relationships because of variations in the demographic composition of subjects across studies (see Brown, 1969; Spence, *et. al.*, 1970; Cunningham, 1967a; and Hasty, 1969). Thus, we need "richer" research designs to simultaneously address these variables.

#### REFERENCES

- Andreasen, A. Attitudes and customer behavior: a decision model. In L. E. Preston (Ed.), New research in marketing. Berkeley: University of California Institute of Business and Economic Research, 1965. Pp. 1-16.
- Arndt, J. Word-of-mouth advertising and informal communication. In D. F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. Pp. 188-239. (a)
- Arndt, J. Perceived risk, sociometric integration, and word-of-mouth in the adoption of a new food product. In D. F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. Pp. 289-316. (b)
- Arndt, J. Role of product-related conversations in diffusion of a new product. Journal of Marketing Research, 1967, 4, 291-295. (c)
- Arndt, J. New product diffusion: the interplay of innovativeness, opinion leadership, learning, perceived risk and product characteristics. Markeds Kommunikasjon, 1968, 2, 1-9. (a)
- Arndt, J. A test of the two-step flow in diffusion of a new product. Journalism Quarterly, 1968, 45, 457-465. (b)
- Arndt, J. Profiling consumer innovators. In J. Arndt (Ed.), Insights into consumer behavior. Boston: Allyn and Bacon, 1968. Pp. 71-83. (c)
- Arndt, J. Intrafamilial homogeneity for perceived risk and opinion leadership. Journal of Advertising, 1972, 1, 40-47.
- Barach, J. A. Advertising effectiveness and risk in the consumer decision process. Journal of Marketing Research, 1969, 6, 314-320.

- Bauer, R. A. Consumer behavior as risk-taking. In R. S. Hancock (Ed.), Dynamic marketing for a changing world. Chicago: American Marketing Association, 1960. Pp. 389-398. Cited from D. F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. Pp. 23-33.
- Bauer, R. A. Risk handling in drug adoption: Role of company preferences. Public Opinion Quarterly, 1961, 25, 546-549.
- Bauer, R. A. Source effect and persuasibility: A new look. In D. F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. Pp. 559-578.
- Bauer, R. A., and Wortzel, L. H. Doctor's choice: The physician and his sources of information about drugs. Journal of Marketing Research, 1966, 3, 40-47.
- Berlyne, D. E. Structure and direction in thinking. New York: Wiley, 1965.
- Bettman, J. R. Perceived risk: A measurement methodology and preliminary findings. In M. Venkatesan (Ed.), Proceedings, Third Annual Convention of the Association for Consumer Research, 1972, 394-403.
- Bettman, J. R. Perceived risk and its components: A model and empirical test. Journal of Marketing Research, 1973, 10, 184-189.
- Bettman, J. R. Information integration in perceived risk. University of California, Los Angeles, Center for Marketing Studies. Working paper number 15, May, 1974.
- Brown, H. E. The role of personality and perceived risk in the purchase of branded headache and pain and cold remedies. Unpublished doctoral dissertation, Ohio State University, 1969.
- Bursk, E. Consumer behavior research seminar manuscript. Boston: Harvard Business School, 1960.
- Coleman, J., Katz, E., and Menzel, H. The diffusion of an innovation among physicians. Sociometry, 1957, 20, 253-269.
- Coleman, J., Menzel, H., and Katz, E. Social pressures in physicians' adoption of a new drug. Journal of Chronic Diseases, 1959, 9, 1-19.
- Copley, T. P. and Callum, F. L. Industrial search behavior and perceived risk. In D. M. Gardner (Ed.), Proceedings, Second Annual Convention of the Association for Consumer Research, 1971, 208-217.
- Cox, D. F. Risk handling in consumer behavior -- an intensive study of two cases. In D. F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. Pp. 34-81. (a)
- Cox, D. F. The influence of cognitive needs and styles on information handling in making product evaluations. In D. F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. Pp. 370-393. (b)

- Cox, D.F. Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. (c)
- Cox, D.F. and Rich, S.V. Perceived risk and consumer decision-making -- the case of telephone shopping. Journal of Marketing Research, 1964, 1, 32-39.
- Cunningham, S.M. Perceived risk as a factor in product-oriented word-of-mouth behavior: A first step. In L.G. Smith (Ed.), Reflections on Progress in Marketing. Chicago: American Marketing Association, 1964. Pp. 229-238.
- Cunningham, S.M. Perceived risk as a factor in the diffusion of new product information. In R.M. Hass (Ed.), Science, Technology, and Marketing. Chicago: American Marketing Association, 1966. Pp. 698-721.
- Cunningham, S.M. The major dimensions of perceived risk. In D.F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. Pp. 82-108. (a)
- Cunningham, S.M. Perceived risk as a factor in informal consumer communications. In D.F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. Pp. 265-288. (b)
- Cunningham, S.M. Perceived risk and brand loyalty. In D.F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. Pp. 507-523. (c)
- Deering Barbara J. and Jacoby, J. Risk enhancement and risk reduction as strategies for handling perceived risk. In M. Venkatesan (Ed.), Proceedings, Third Annual Convention of the Association for Consumer Research, 1972, 404-416.
- Engel, J.F., Kollat, D.T., and Blackwell, R.D. Consumer behavior. (2nd ed.) New York: Holt, Rinehart, Winston, 1973.
- Green, P.E. Consumer use of information. In J.W. Newman (Ed.), On knowing the consumer. New York: John Wiley and Sons, 1966. Pp. 67-80.
- Hart, E.W., Jr. Consumer risk-taking for self and for spouse. Unpublished doctoral dissertation, Purdue University, 1974.
- Hasty, R.W. Some demographic correlates of risk perception in the purchase of electric housewares. Unpublished doctoral dissertation, University of Colorado, 1969.
- Hisrich, R.D., Dornoff, R.J., and Kernan, J.B. Perceived risk in store selection. Journal of Marketing Research, 1972, 9, 435-439.
- Howard, J.A. and Sheth, J.N. The theory of buyer behavior. New York: John Wiley and Sons, 1969.
- Jacoby, J. and Kaplan, L. The components of perceived risk. In M. Venkatesan (Ed.), Proceedings, Third Annual Convention of the Association for Consumer Research, 1972, 382-393.
- Kaplan, L., Szybillo, G.J., and Jacoby, J. Components of perceived risk in product purchase: A cross-validation. Journal of Applied Psychology, 1974, 59, 287-291.

- Levitt, T. Persuasibility of purchasing agents and chemists: Effects of source, presentation, risk, audience competence, and time. In D. F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. Pp. 541-558.
- Lutz, R. J. and Reilly, P. J. An exploration of the effects of perceived social and performance risk on consumer information acquisition. In S. Ward and P. Wright (Eds.), Advances in Consumer Research. Volume 1. Proceedings of the Fourth Annual Convention of the Association for Consumer Research, 1973, 393-405.
- McMillan, J. R. The role of perceived risk in vendor selection decisions. Unpublished doctoral dissertation, Ohio State University, 1972.
- Newton, D. A. A marketing communications model for sales management. In D. F. Cox (Ed.), Risk-taking and information-handling in consumer behavior. Boston: Harvard University Press, 1967. Pp. 579-602.
- Nicosia, F. M. Perceived risk, information processing, and consumer behavior: A review article. Journal of Business, 1969, 42, 162-166.
- Perry, M. and Hamm, B. C. Canonical analysis of relations between socio-economic risk and personal influence in purchase decisions. Journal of Marketing Research, 1969, 6, 351-354.
- Pettigrew, T. F. The measurement and correlates of category width as a cognitive variable. Journal of Personality, 1956, 26, 532-544.
- Popielarz, D. T. An exploration of perceived risk and willingness to try new products. Journal of Marketing Research, 1967, 9, 368-372.
- Reingen, P. H. Choice shifts along a dimension of risk tested with consumer-related stimuli. Unpublished doctoral dissertation, University of Cincinnati, 1974.
- Roselius, T. Consumer rankings of risk reduction methods. Journal of Marketing, 1971, 35, 56-61.
- Ross, I. Structure of information seeking and personality in consumer decision-making. Journal of Business Administration, 1972, 3, 55-67.
- Schiffman, L. G. Perceived risk in new product trial by elderly consumers. Journal of Marketing Research, 1972, 9, 106-108.
- Sheth, J. N. Perceived risk and diffusion of innovations. In J. Arndt (Ed.), Insights into consumer behavior. Boston: Allyn and Bacon; 1968. Pp. 173-188.
- Sheth, J. N. and Venkatesan, M. Risk-reduction processes in repetitive consumer behavior. Journal of Marketing Research, 1968, 5, 307-310.
- Slovic, P. and Lichtenstein, Sarah. Relative importance of probabilities and payoff in risk taking. Journal of Experimental Psychology Monographs, 1968, 78, 1-18.
- Spence, H. E., Engel, J. F., and Blackwell, R. D. Perceived risk in mail-order and retail store buying. Journal of Marketing Research, 1970, 7, 364-369.

- Taylor, J. W. The role of risk in consumer behavior. Journal of Marketing, 1974, 38, 54-60.
- Venkatesan, M. Cognitive consistency and novelty-seeking. In S. Ward and T. S. Robertson (Eds.), Consumer behavior: Theoretical sources. New York: Prentice Hall, 1973.
- Wilding, J. and Bauer, R. A. Consumer goals and reactions to a communication source. Journal of Marketing Research, 1968, 5, 73-77.
- Zikmund, W. G. and Scott, J. E. A multivariate analysis of perceived risk, self-confidence and information sources. In S. Ward and P. Wright (Eds.), Advances in Consumer Research. Volume 1. Proceedings of the Fourth Annual Convention of the Association for Consumer Research , 1973, 406-416.





COGNITIVE DISSONANCE AND CONSUMER BEHAVIOR: A REVIEW  
OF THE EVIDENCE

William H. Cummings and M. Venkatesan<sup>1</sup>  
University of Iowa

The theory of cognitive dissonance has generated a good deal of research in consumer behavior; however, this research has not yet been fully reviewed and critiqued. The relevant evidence from three phases of consumer behavior is reviewed here: pre-decisional determinants of product preference, post-decisional determinants of product preference, and information seeking behavior. The evidence from the first two areas generally supports the dissonance-based predictions, while the evidence from the third area generally fails to support the predictions.

There have been a number of nasty rumors lately that "cognitive dissonance theory is dead" and that "cognitive dissonance theory began to be out of date a long time ago." This may well be so. It would seem, however, that before we bury dissonance theory, we should make sure that all the phenomena and events which it explains (particularly post-decisional attitude change) are also dead or out of date. We should make sure that the theory is entirely incorporated within new and more general theories. And we should be sure that dissonance theory is incapable of generating further predictions of interest to social scientists.

A few years ago, dissonance theory generated considerable interest and enthusiasm. The generalizability of the theory from the social psychology laboratory to consumer behavior appeared quite certain. Twenty-three studies have since examined the applicability of dissonance theory to consumer behavior.<sup>2</sup> Although a number of reviews have examined the theoretical relevance of dissonance theory to the consumer behavior area, no review has apparently yet attempted to integrate and critique the empirical evidence on this matter. This will be the major purpose of this review. Before we lay dissonance theory to rest, it may be worthwhile to examine the studies and answer a few questions:

- (a) What does the available evidence tell us about the applicability of dissonance theory (and related approaches) to consumer behavior?
- (b) Does the evidence (both positive and negative) stand up to methodological scrutiny? Can the methodological problems be avoided in the future?
- (c) What sort of research is needed for a better understanding of the applicability--if any--of dissonance theory (and related approaches) to consumer behavior?
- (d) What interesting predictions--if any--can still be derived from dissonance theory and related approaches?

The initial presentation of the theory of cognitive dissonance (Festinger, 1957) clearly demonstrated the relevance of dissonance theory to the purchase situation. Festinger derived his theory from two basic principles: (a) Dissonance is uncomfortable and will motivate the person to reduce it, and (b) the dissonant individual will avoid situations which produce

further dissonance. Dissonance is a post-decisional phenomenon and should therefore be a post-purchase phenomenon. Dissonance theory works according to a reality principle in that the cognition least anchored in reality will be the one most likely to change; since the purchase will be a strongly anchored behavior, the cognitions most likely to change will be those concerning satisfaction with the purchased product relative to rejected products. Dissonance should increase with the importance of the cognitions, and to the extent that the individual commits some time and some money in the purchase, many purchase decisions should be important ones. So the potential application of dissonance theory to consumer behavior is considerable. In answering the question of whether the empirical findings have supported this application, we have found it convenient to divide the research into three areas: (a) pre-decisional determinants of product preference/satisfaction; i.e., those variables which increase magnitude of dissonance; (b) post-decisional determinants of product preference/satisfaction; i.e., those variables which decrease magnitude of dissonance; and (c) effects of dissonance arousal on information seeking behavior. We will now examine the evidence within each area.

#### Pre-decisional Determinants of Product Preference

Most applications of dissonance theory in consumer behavior fall into this category--attempts to modify product satisfaction via the control of events prior to the purchase decision. Brehm provided the landmark study in 1956 on the effect of magnitude of dissonance arousal on attitude change. Brehm observed--as predicted by dissonance theory--that when subjects had to choose between two products which were similar in desirability, there was substantial attitude change in favor of the chosen product relative to the unchosen product. When subjects chose between two products that were disparate in desirability, there was much less attitude change. In all conditions, the chosen product was favorably reevaluated and the unchosen product was derogated or unchanged in desirability.

Later studies have replicated and extended Brehm's findings. LoSciuto & Perloff (1967) observed significantly more favorable reevaluation of the chosen product (relative to the unchosen product) when subjects had to choose between two record albums close in desirability than when subjects chose between two albums disparate in desirability. Thus, the more similar the products which are considered, the greater is the magnitude of dissonance, and the greater the "spread" in ultimate preference rankings of the two products. Sheth (1968, 1970) also demonstrated the predicted positive reevaluation of chosen products and derogation of unchosen products for various product classes, for both students and housewives.

Likewise, Anderson, et al. (1966), have demonstrated the predicted post-decisional spread in desirability of the alternatives. Anderson, et al., also demonstrated that the reevaluation tendency increased with the number of alternatives. However, this effect was significant only when the initial difference in desirability between the alternatives was large. The number of alternatives seems to have little effect on magnitude of dissonance if dissonance is already quite high.

Mittelstaedt (1969) used a modified free choice paradigm to demonstrate that "buyers" who became committed to one particular "brand" would persist in their preference for that "brand," even when offered the choice of a "brand" they originally preferred more. Subjects first rank-ordered nine

swimsuits and were then given a choice either between their third-ranked and fourth-ranked choices (high dissonance condition) or between their third-ranked and fifth-ranked choices (low dissonance condition). Later, most of the high dissonance "purchasers" preferred the chosen (third-ranked) swimsuit over a swimsuit they originally ranked higher. Most of the low dissonance subjects did not.

A simulated purchase study by Holloway (1967) examined the effects of inducement to purchase, cognitive overlap, and presence/absence of supportive information. These factors had the predicted--but a nonsignificant--effect on product reevaluations. Holloway also manipulated subjects' anticipated dissonance and found (contrary to expectations) greater product reevaluation under low anticipated dissonance, but, again, nonsignificantly so.

Dissonance theory also predicts that product satisfaction should increase when the amount of effort expended in obtaining the product is increased. Cardozo (1965) found that subjects who expended a good deal of effort to obtain a product (a ball-point pen) rated this product significantly higher than did subjects who expended little effort. However, this effect was significant only when subjects were expecting a higher quality product than they actually received.

A common feature of all the above studies of consumer behavior with one exception (Sheth, 1970) is that they were conducted in the laboratory. Doob, et al., (1969) have employed a field study and a basic prediction from the forced compliance paradigm, e.g., "the less the pressure...put upon the person to perform the act, the greater the dissonance (Kiesler et al., 1969, p. 206)." Doob, et al., observed that introducing a product brand at a special, low, discount price would hurt later sales of that product. Thus, the higher the initial price paid for the product, the greater the magnitude of dissonance aroused, the greater the ultimate satisfaction should be, and thus, the greater the tendency to repurchase that brand. Doob, et al., matched this prediction for five different household product classes.

The findings presented above are summarized in Table 1. This summary indicates that 11 findings have supported predictions from dissonance theory, while 2 findings have failed to support the predictions, and there are 4 findings which could be considered as "mixed." Before drawing any conclusions from this simple count, we must examine the findings in view of the methodological problems in this area. The methodological problems of each study are presented in the third column of Table 1.

One criticism that has been generally made of investigations in consumer behavior which have employed the free choice paradigm is the criticism of "regression artifact" or "ceiling effect." Specifically, Oshikawa (1968) has demonstrated that free choice-type reevaluations made in the absence of any choice may show a dissonance-type effect. But when corrected for ceiling effect, this dissonance-type effect disappears. To be sure, however, two of the five free choice studies reported above have included a correction or control for this sort of regression artifact (Brehm, 1956; LoSciuto & Perloff, 1967). In one other study (Sheth, 1970), findings were reported that directly contradict the ceiling effect (see Sheth, 1971). Thus, the appropriate control or correction for regression should be included in any study where regression artifact is a possibility. However, these studies have, in general, demonstrated the predicted reranking of the chosen and unchosen products, despite the possibility of regression artifact.

TABLE 1

## Pre-decisional Determinants of Product Preference

Study	Major Findings	Major Problems
1. Brehm (1956)	Pos. reevaluation of chosen product; Neg. (or no change) reevaluation of unchosen product; greater spread in desirability following close choice; greater spread (sig. only) when additional info. was supplied.	Rejection of large number of subjects (see Chapanis and Chapanis, 1964).
2. Anderson, et al. (1966)	Pos. reevaluation of chosen product relative to unchosen; greater spread following close choice; greater spread with four than with two alternatives (sig. only under disparate choice).	Possible regression artifact (ceiling effect).
3. LoSciuto and Perloff (1967)	Greater pos. reevaluation of chosen product relative to unchosen following close choice; effect maintained at least one week following choice.	
4. Sheth (1968, 1970)	Pos. reevaluation of chosen products; Neg. reevaluation of unchosen products; greater spread in desirability following close choice.	Possible regression artifacts.
5. Mittelstaedt (1969)	Greater pos. reevaluation of chosen product following close choice (implications for brand loyalty).	Role-playing technique.
6. Holloway (1967)	Nonsignificant effects of inducement, cognitive overlap, presence of supportive info., and anticipated dissonance.	Role-playing technique.
7. Cardozo (1965)	Evaluation of obtained product increased with effort; effect greatest (sig. only) under high expectation.	Prerequisite conditions (volition and importance) absent.
8. Doob, et al. (1969)	Positive inducement impairs later sales of product.	Compelling alternative explanations.
9. Cohen and Houston (1972)	Positive relation between brand loyalty and perceived quality of chosen product.	Correlational evidence.
10. Cohen and Goldberg (1970)	No relation between pre-decisional conflict and subsequent measures of product preference.	Correlational evidence; questionable relevance of conflict measure to dissonance theory (see Oshikawa, 1972; Hawkins, 1972).

Another problem of studies in this area is a possible failure to meet the prerequisite conditions for producing dissonance. Accumulated evidence in social psychology has suggested that predictions from dissonance theory cannot clearly be made unless the choice situation meets three conditions: (a) volition concerning the choice, (b) irrevocable commitment to the decision (product choice), and (c) importance of the choice to the individual's self-concept. The prerequisite of irrevocable commitment has generally been met, while the prerequisite of volition has not always been met. The condition of importance of choice is most questionable: Has it always been upheld? Certainly the role-playing situations, which have been employed in some studies, would necessarily reduce the importance of the decision for the individual and his future behavior. Although some studies which have failed to clearly meet the prerequisite conditions may have supported the theory, these effects may be due to causes other than dissonance arousal.

Some studies which have cited dissonance theory as a source of hypotheses do not, in fact, appear to be relevant to the main theory. For example, three studies have examined the effect of prior expectations on subsequent product evaluation (Cardozo, 1965; Olshavsky and Miller, 1972; Anderson, 1973). While this is certainly an important question, it is a question distantly relevant to dissonance theory. The effects of such manipulations cannot be taken as either support or failure to support the theory.

Two studies cited in Table 1 have taken correlational findings as causal support or disconfirmation of the theory. For example, one study found no effects of subjects' post-decisional ratings of pre-decisional conflict on subsequent measures of brand preference, intention to repurchase, or brand loyalty. These findings were interpreted as a disconfirmation of dissonance theory. What the results do demonstrate is that there is little if any correlation between the initial measure of conflict and the subsequent measures. Moreover, the technique of measuring dissonance by asking the subject about his level of conflict or worry has been questioned (Oshikawa, 1972; Hawkins, 1972). These investigations have demonstrated that such measures are correlated with measures of general confidence and anxiety.

A final problem in interpreting these studies is the compelling alternative explanations generated by studies which appear to be otherwise methodologically sound. For example, Doob et al. (1969) demonstrated that introducing a product at a special low price decreases later sales, as predicted by the theory. But there was no safeguard against buyers stockpiling goods at the initial purchase, during the sale. Also, there was no indication that a sale was in progress, and some buyers may have been "frustrated" after the sale when they were suddenly faced with a more expensive product.

In summary, the evidence has generally supported the predicted relationships between pre-decisional factors and subsequent product preference. No finding from this area has conclusively disconfirmed any prediction from the theory. On the other hand the positive evidence is importantly qualified by these major methodological concerns. But it should be noted that the evidence in favor of the applicability of dissonance theory to this area is more voluminous and somewhat more substantial than the evidence against. There is no single methodological or conceptual limitation that runs throughout all the positive instances of the dissonance effect. There is no other single explanation -- other than that of cognitive dissonance theory -- that can fully account for the results of these studies. These findings, then, have substantiated the predicted relation between pre-decisional factors which would tend to affect magnitude of dissonance and subsequent product preference.

### Post-decisional Determinants of Product Preference

Three studies have employed a postpurchase reinforcement technique to increase purchasers' satisfaction with the obtained product. Two studies found greater product satisfaction with the technique, as predicted, while the third study obtained mixed results. Table 2 summarizes these findings. VanDyke (1966) demonstrated that a supportive letter from the dealer increased consonance<sup>3</sup> following the purchase of a new car. Purchasers who did not receive the supportive letter showed no such increase in consonance. Donnelly and Ivancevich (1970) found that two supportive phone calls within two weeks following the purchase of a new car would decrease the backout rate from about six percent to about two and one-half percent. Hunt (1970) likewise found that a supportive letter following the purchase of a refrigerator reduced post-transaction anxiety and increased disposition to repurchase. However, a supportive telephone contact -- which was similar to the letter in content -- had the opposite effect. (Both effects were statistically nonsignificant.)

TABLE 2

#### Post-decisional Determinants of Product Preference

Study	Major Findings	Major Problems
1. VanDyke (1966)	Supportive letter after purchase of car increased consonance.	Compelling alternative explanation via reinforcement theory.
2. Donnelly and Ivancevich (1970)	Supportive phone calls following purchase of car decreased back-out rate.	
3. Hunt (1970)	Supportive letter following purchase of refrigerator decreased post-transactional anxiety; supportive phone call increased anxiety.	Questionable measures of dissonance reduction (e.g., "conflict").

How shall we interpret these results? First, it is obvious that the positive effects (postpurchase contact increasing satisfaction) are all interpretable within common reinforcement notions. Postpurchase reinforcement provided in the form of social support should directly increase produce satisfaction and tendency to repurchase. This is a much simpler explanation than is dissonance theory. The interpretation of these studies via dissonance theory, however, is beneficial for two reasons. (a) Dissonance theory explains why these social contacts should be particularly reinforcing immediately following the purchase, i.e., because the purchaser of a major product is experiencing considerable dissonance and is in need of social support. (b) The dissonance interpretation is consistent with the interpretation given the studies in the first section (predecisional determinants of product satisfaction). Thus, while dissonance theory can explain both sets of findings (pre and post), reinforcement notions work best only with the latter set of findings.

Thus, despite the presence of one negative finding and some compelling alternative explanations, it appears that dissonance theory offers the most parsimonious account of the studies reviewed in this section.

## Effects of Dissonance Arousal on Information Seeking Behavior

Festinger's second basic hypothesis regarding dissonance theory was that the dissonant individual should actively avoid information which would tend to increase the dissonance and seek information which would tend to support his decision. However, the evidence within social psychology -- obtained under optimal laboratory conditions -- has failed to support this hypothesis. Freedman and Sears (1965) and Sears (1968) have concluded that the available evidence does not support either (a) a general preference for supportive over nonsupportive information or (b) a greater information seeking/avoidance tendency by high dissonance subjects.

Is there evidence within the area of consumer behavior to modify this conclusion? The relevant findings are summarized in Table 3, and the answer is a clear "no," there is no evidence here which strongly supports a link between magnitude of dissonance and information seeking behavior. All the findings in this area are, in turn, qualified by some compelling methodological problems. Ehrlich et al. (1957) and Engel (1963) found the predicted effects of magnitude of dissonance on some measures but failed to obtain the predicted effects on a number of additional measures. Kassarian and Cohen (1965) found the predicted inverse relation between the amount a person smoked and how believable that person found the smoking-cancer link to be. However, this is correlational evidence. Perhaps strongly anchored smoking behavior reduced belief in the cancer evidence. Or perhaps people who didn't believe the link became the heaviest smokers. Or maybe a certain type of person tended to be both a nonbeliever and a heavy smoker. LoSciuto and Perloff (1967) found no effect of magnitude of dissonance on a recognition task. However, this recognition task followed a reranking task on which differential reranking of the products was obtained. If dissonance was reduced by reranking the products, it is unlikely that there will be any effect of magnitude of dissonance on subsequent trials.

Thus, the evidence on this issue is most equivocal. In view of repeated failures to demonstrate any consistent effect of magnitude of dissonance on information seeking, avoidance, and perceptual processing, we cannot conclude at this time that dissonance arousal factors are relevant to postpurchase information seeking variables in consumer behavior.

### Summary, Conclusions, and Extensions

We feel that the relevant evidence on predecisional and postdecisional determinants of product preference (satisfaction) has given generous support to dissonance theory. Each study reviewed here has been interpreted in terms of its methodological and conceptual limitations. In view of these limitations, the negative evidence has not conclusively disconfirmed any aspect of the theory. Moreover, dissonance theory appears to offer a unique and parsimonious explanation for the positive findings. We conclude, then, that the generalizability of cognitive dissonance theory to consumer behavior -- excluding information seeking behavior -- has been substantiated.

The questions which would now appear to be of primary interest in this area are the interaction questions; specifically, what are the precise conditions under which we would expect the dissonance effects to occur. The studies in the consumer behavior area have not been primarily oriented toward this question. From the social psychological literature we know that predictions from the theory can be most accurately applied when there is high



TABLE 3

## Effects of Dissonance Arousal on Information Seeking Behavior

Study	Major Findings	Major Problems
1. Ehrlich et al. (1957)	Recent car purchasers (high diss) read more consonant than dissonant ads; recent purchasers read more consonant ads than did old car owners (low diss); no effect of recency of purchase (diss) on reading of dissonant ads, recall of ad features, and choice to read new ads.	Rejection of large number of <u>Ss</u> ; lack of a proper control condition (see Chapanis and Chapanis, 1964); failure to shield against alternative modes of diss reduction.
2. Engel (1963)	Recent car purchasers (high diss) showed greater recall of consonant ads than did nonowners (low diss); no effects of diss on recognition, agreement with cons./diss. statements; memory of cons./diss. statements.	Failure to shield against alternative modes of diss reduction (see Straits, 1964).
3. Kassarian and Cohen (1965)	Inverse relation between smoking behavior and believability of the smoking-cancer link; rationalization of smoking, deemphasis of health-related aspects of smoking, etc.	Correlational evidence.
4. LoSciuto and Perloff (1967)	No effect of magnitude of dissonance on a recognition task.	Failure to shield against alternative modes of dissonance reduction.

perceived volition concerning the choice, when there is an irrevocable commitment to the choice, and when the choice is an important one (ego-involving) for the individual. These factors raise a number of interesting questions in application: What sorts of products are sufficiently important to arouse dissonance? Are certain product classes more susceptible to dissonance effects than others, even when cost is held constant? What is the optimal balance between volition and positive inducement in maximizing both initial purchase tendency and postpurchase satisfaction? What modes of dissonance reduction are normally used by the consumer in everyday life? What are some of the long-term effects of differential dissonance arousal and reduction; i.e., is the increased product satisfaction that results from dissonance reduction a lasting effect?

Experimental research in this area has concentrated primarily on the free choice paradigm. However, there are other paradigms which may be used for research and application. The following appear promising: (a) forced compliance paradigm, (b) fait accompli paradigm, (c) attribution theory, and (d) commitment theory.

In conclusion, we feel that the application of dissonance theory to consumer behavior -- at least within the areas of predecisional and postdeci-

sional determinants of product satisfaction -- has been substantiated. There have been a number of methodological problems in this line of research, but most of them are of a sort that can be avoided by means of a simple inventory of possible pitfalls or the addition of the appropriate control conditions. It does not appear that dissonance theory can be adequately incorporated within any other framework and retain its power to generate new and intriguing hypotheses. And that power -- to generate hypotheses -- should continue to be the major selling point of dissonance theory.

#### FOOTNOTES

1. William H. Cummings is a doctoral student in social psychology and M. Venkatesan is Professor of Business Administration at the College of Business Administration, The University of Iowa.
2. We have attempted to include all relevant studies from the Consumer Behavior literature while including as few studies as possible from the social psychological literature. Only three studies from the social psychological literature have been included in this review. Three studies from the consumer behavior literature have been excluded. Oshikawa's study (1970) is not directly relevant to consumer behavior. Reizenstein's study (1971) hints at relevance to dissonance theory but fails to provide necessary choice-making conditions for producing dissonance among the subjects. Bell's study (1967) is questionable on basic conceptual grounds, viz., "direct" indication of magnitude of dissonance arousal.
3. VanDyke's estimate of consonance is theoretically appropriate and should be of some interest. The relative weighting of each of the five most important attributes involved in a car purchase (as rated by each subject) was multiplied by the difference between the subject's rating of the chosen brand and the unchosen brand on that attribute. This product was summed over all five attributes and constituted the consonance estimate for each subject.

#### REFERENCES

- Anderson, L. K., J. R. Taylor, and R. J. Holloway. The consumer and his alternatives: An experimental approach. Journal of Marketing Research, 1966, 3, 62-67.
- Anderson, R. E. Consumer dissatisfaction: The effect of disconfirmed expectancy on perceived product performance. Journal of Marketing Research, 1973, 10, 38-44.
- Bell, G. D. The automobile buyer after the purchase. Journal of Marketing, 1967, 31 (3), 12-16.
- Brehm, J. W. Post-decision changes in the desirability of alternatives. Journal of Abnormal and Social Psychology, 1956, 52, 384-389.
- Cardozo, R. N. An experimental study of customer effort, expectation, and satisfaction. Journal of Marketing Research, 1965, 2, 244-249.

- Chapanis, N., and A. Chapanis. Cognitive dissonance five years later. Psychological Bulletin, 1964, 61, 1-22.
- Cohen, J. B., and M. E. Goldberg. The dissonance model in post-decision product evaluation. Journal of Marketing Research, 1970, 7, 315-321.
- Cohen J. B., and M. J. Houston. Cognitive consequences of brand loyalty. Journal of Marketing Research, 1972, 9, 97-99.
- Donnelly, J. H., and J. M. Ivancevich. Post-purchase reinforcement and back-out behavior. Journal of Marketing Research, 1970, 7, 399-400.
- Doob, A. N., J. M. Carlsmith, J. Freedman, T. K. Landauer, and S. Tom. Effect of initial selling price on subsequent sales. Journal of Personality and Social Psychology, 1969, 11, 345-350.
- Ehrlich, D., I. Guttman, P. Schonbach, and J. Mills. Postdecision exposure to relevant information. Journal of Abnormal and Social Psychology, 1957, 54, 98-102.
- Engel, J. F. Are automobile purchasers dissonant consumers? Journal of Marketing, 1963, 27, 55-58.
- Festinger, L. A theory of cognitive dissonance. Stanford: Stanford University Press, 1957.
- Freedman, J. L., and D. O. Sears. Selective exposure. In L. Berkowitz (Ed.), Advances in experimental social psychology. Vol. 2. New York: Academic Press, 1965.
- Hawkins, D. I. Reported cognitive dissonance and anxiety: Some additional findings. Journal of Marketing, 1972, 36 (3), 63-66.
- Holloway, R. J. An experiment on consumer dissonance. Journal of Marketing, 1967, 31 (1), 39-43.
- Hunt, S. D. Post-transaction communications and dissonance reduction. Journal of Marketing, 1970, 34 (3), 46-51.
- Kassarjian, H. H., and J. B. Cohen. Cognitive dissonance and consumer behavior. California Management Review, 1965, 8, 55-64.
- Kiesler, C. A., B. E. Collins, and N. Miller. Attitude change: A critical analysis of theoretical approaches. New York: Wiley, 1969.
- LoSciuto, L., and R. Perloff. Influence of product preference on dissonance reduction. Journal of Marketing Research, 1967, 4, 286-290.
- Mittelstaedt, R. A dissonance approach to repeat purchasing behavior. Journal of Marketing Research, 1969, 6, 444-446.
- Olshavsky, R. W., and J. A. Miller. Consumer expectations, product performance, and Perceived product quality. Journal of Marketing Research, 1972, 9, 19-21.

- Oshikawa, S. The theory of cognitive dissonance and experimental research. Journal of Marketing Research, 1968, 5, 429-430.
- Oshikawa, S. Can cognitive dissonance theory explain consumer behavior? Journal of Marketing, 1969, 33 (4), 44-49.
- Oshikawa, S. Consumer pre-decision conflict and post-decision dissonance. Behavioral Science, 1970, 15, 132-140.
- Oshikawa, S. Dissonance reduction or artifact? Journal of Marketing Research, 1971, 8, 514-515.
- Oshikawa, S. The measurement of cognitive dissonance: Some experimental findings. Journal of Marketing, 1972, 36 (4), 64-67.
- Reizenstein, R. C. A dissonance approach to measuring the effectiveness of two personal selling techniques through decision reversal. Combined Proceedings, American Marketing Association, Spring and Fall, 1971, 176-180.
- Sears, D. O. The paradox of de facto selective exposure without preference for supportive information. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, and P. H. Tannenbaum (Eds.), Theories of cognitive consistency: A sourcebook. Chicago: Rand-McNally, 1968.
- Sheth, J. N. Cognitive dissonance, brand preference, and product familiarity. In J. Arndt (Ed.), Insights into consumer behavior. Boston: Allyn and Bacon, 1968.
- Sheth, J. N. Are there differences in dissonance reduction behavior between students and housewives? Journal of Marketing Research, 1970, 7, 243-245.
- Sheth, J. N. Dissonance reduction or artifact? A reply. Journal of Marketing Research, 1971, 8, 516-517.
- Straits, B. C. The pursuit of the dissonant consumer. Journal of Marketing, 1964, 28, 62-66.
- VanDyke, J. E. A study of the arousal and reduction of cost-decision dissonance following the purchase of a new automobile. Unpublished master's thesis, The University of Iowa, 1966.



A MODERN SOCIOLOGICAL APPROACH TO THE STRATIFICATION OF  
MATERIAL LIFE STYLES

Marcus Felson  
University of Illinois at Urbana

After discussing the fallacies of the Warnerian model of social class, the author criticizes modern marketing research for its lack of adequate theoretical models. Special criticism is reserved for "life style" research, a current fad which fails to specify solid models of consumer decisions. An alternative approach is an opportunity model, taking time and knowledge as resources in addition to income. The causal impact of one consumer decision on another must be formulated and precise models of how social opportunities affect consumption must be specified.

One of the least successful areas of social research is the sociological study of material life styles. Many market researchers are inclined to neglect modern sociology, while modern sociologists generally neglect the study of consumer taste. The present paper is designed to bring market researchers up to date on developments in stratification research over the last decade and to suggest how they might take advantage of these developments in planning their own research. Recent literature reviews are available elsewhere (cf. Lasswell and Benbrook, 1974; Dietrick, 1974). My purpose here is to summarize the state of the art and to suggest applying its research tools to studying the stratification of consumers.

What most market researchers call "social class" is a 20-year-out-of-date concept popularized by W. Lloyd Warner. The Warnerian Model of social class is characterized by (1) reasonably clear-cut class boundaries, (2) a single dimension summarizing social inequality, and (3) a single type of relationship between social class and other variables.

Warner's model simply cannot stand up to the evidence collected in modern America for several reasons. First, the correlations among stratification variables taking individuals or households as units of analysis are not so high that one can reasonably treat social class as a single dimension. Often these correlations are .4 or .3, seldom over .6. Second, stratification variables have causal relationships to one another which Warner ignored. Education has an effect on occupational attainment which has an effect on income. Yet Warner mixed all these distinct variables together, ignoring their causal relationships. Third, education, occupation, income, race, and other stratification variables often have different sorts of consequences for social life, consequences which have been decomposed by more careful researchers. Fourth, Warner failed to consider how various rewards come at various stages of the life cycle or even change in the period of a few years. Poverty, prosperity and occupation are conditions rather than properties of people, so that they are not demographic variables at all. Being an electrician at Company X is not an innate or immutable characteristic, so that treating it as "demographic" completely misses the point. regrettably, this bad habit has been adopted by modern market research. Fifth, Warner did not realize that

clear boundaries between social strata simply do not exist with respect to most stratification variables. The terms "blue collar," "white collar" and "middle class," however convenient, obscure the actual distributions of occupational prestige and income, which are actually smooth curves with few clear fissures. Education does have clearer gaps because diplomas are offered mainly after the 12th and 16th year, but even this so-called gap is mitigated by the fact that each year of education has a payoff regardless of whether diplomas are received. The use of class boundaries or class names is largely arbitrary and their reification is a good example of careless inference from data. The major exception has to do with race, which tends to be a dichotomous variable with virtually no one holding a position intermediate between black and white.

Modern stratification models generally decompose "socioeconomic status" into components which are then related in a causal model. Assumptions about cause and effect should be justified before setting out and intervening variables should be precisely specified. Indirect effects should be allowed, and variables having no direct effect upon one another are welcome for consideration. Blind factor analysis or any computerized searching of correlation matrices before specifying a causal model is strongly discouraged in modern stratification research. Market research will need to follow the same path, not only in relating social stratification to consumption but also in relating consumer variables to one another. I will come back to the latter point, but first I wish to comment upon a tendency to try to maximize the proportion of variance explained. Under the leadership of Otis Dudley Duncan sociologists have largely escaped from that quicksand (see Blau and Duncan, 1967, as an example). It does science little good to have a high correlation if you have a spurious causal model explaining that correlation. Of course, raw and spurious correlations can be used by businessmen to decide in which magazine to advertise their product. But any creative endeavor to develop new products cannot proceed without understanding. Furthermore, marketing as the science of consumer taste needs to go beyond selling products and to build a model of consumption in a larger sense. The Williamsberg Conference on Social Structure, Family Life Styles, and Economic Behavior (see Sheldon, 1973) took a major step in this direction, but marketing as a discipline needs to carry out the task. I do not deny that plenty of very general schematic diagrams are found in marketing journals and books, but these do little good if they are ignored when data is collected and analyzed. Causal modeling, as done by the Duncan School and its followers in sociology, forces researchers to operationalize their theoretical statements. If it is later discovered that lots of variance cannot be explained, this may reveal that non-systematic or random factors influence behavior. The degree of freedom in a family budget and the thousands of goods and services available tend to reduce the correlations between any one consumer trait and its predictors for good substantive reasons. The proper measure relating many types of consumption to their predictors is often an odds ratio or conditional probability, rather than a standardized regression or correlation coefficient. In a sense, any correlation above .1 relating very specific consumption to serious predictors is a miracle. In many cases, the small regression coefficients linking independent to dependent variables may arise from an incomplete model. Measurement error may attenuate the relationships. Or the failure to specify intervening variables may give the illusion of no effect. Consider the case where four variables form a causal chain: A affects B which affects C which affects D. Let us say that each of these causal coefficients is a respectable .5, standardized. Now suppose that the researcher omits intervening variables B and C. It will appear that the effect of A on D is  $(.5)(.5)(.5) = .0625$ . Though this is small, it obscures a significant

process which helps one to understand how consumer behavior works. Perhaps the main reason why the multiple  $R^2$  derived from regressing any particular consumption variable against any set of background variables fails to be very large is that these models are misspecified in a very serious way. Consumer variables tend to have negative impacts on one another, insofar as consumer goods and services are either substitutable for one another or compete with one another for scarce resources. It is interesting that consumer research models tend to ignore this basic principle of economics. This produces small  $R^2$  for the simple arithmetic reason that one cannot have high correlations between background variables and consumer traits themselves strongly negative in relationship. Such a situation would lead to correlations greater than 1, as can be discovered by applying the fundamental theorem of path analysis. Similarly, elementary arithmetic will show that competing consumer goods must have rather small correlations with any sort of background variable. Multiple regression analysis based upon such misspecified models will inevitably lead to small coefficients and small  $R^2$ . The conclusion that socioeconomic status and other background variables have little consequence for consumption is an artifact of this misspecified model and totally unjustified. Regression analysis should only be used if the researcher keeps in the back of his or her mind the fact that regression will suppress the true effects of background variables on consumption. Use of more elaborate causal models is preferable, once problems of identifiability can be worked out. In the meantime, conclusions that socioeconomic status is unimportant for consumption are best treated as false.

I would also argue that the relationship between background factors and consumer traits is often obscured by the following suppressor variables:

- (1) Quantity of time available to shop or to consume
- (2) Scheduling of work, school and other obligations which structure use of time
- (3) Knowledge about consumer goods and their attributes
- (4) Life cycle and family variables, carefully partialled out
- (5) Female labor force participation

My general thesis is that consumption is organized by the structure of opportunities available to consumers. Material resources are only one limitation upon opportunities, and the weakness of the relationship of income to many consumer traits is that so many other variables also affect opportunity to consume a given product. Indeed, the rising median level of per capita prosperity over the past half century has decreased the family's dependence upon income, thereby reducing income stratification of life styles (see Felson, 1974). New resources are likely to take on importance. Time is first on the list because a person needs time to buy and to use consumer goods. If time is limited, certain goods facilitate its conservation. Not only is the quantity of available time important but also the nature of one's schedule. A man and wife each of whom works 40 hours a week will consume differently if one of them must work a night shift. How many cars they need, which groceries they purchase, who does the shopping, what activities are shared--all of these are affected by the scheduling of worklife.

Knowledge is a very important and neglected resource. How can someone deliberately purchase a product which he or she has never heard of? How



can a product's high class image sell it to people who think it has a low class image? Much of the so-called unexplained variance in preference for products or brands probably has to do with variance in public knowledge about them. Thus any model of consumption needs to include knowledge or lack of knowledge about consumption. Age, life cycle, family size and other such attributes condition consumer choice by putting people into positions which require or encourage certain behavior patterns. These variables have different effects, depending upon which consumer trait is under study. The failure of many earlier attempts to use these variables as predictors is probably attributable to careless model specifications. In some cases, whether you have any children is more important than how many you have. Sometimes teenage children affect consumption in ways that younger children do not. Age is not a direct measure of life cycle or role obligations, so that its impact on consumption requires consideration of other traits (e.g., marital status). The point is that social roles and statuses affect opportunities to consume various goods, but do so in precise ways which must be disentangled.

Female labor force participation is important because it lessens a woman's free time and alters her exposure to other people. It also may increase her power in her own family. In any case, it affects her opportunities to consume.

None of these suggestions will do a bit of good unless market researchers analyze more fully the opportunity to consume and build a precise model which breaks down family consumption step by step. Poor theory yields poor correlations and even if by chance one finds large correlations, poor theory will render them meaningless.

Present marketing theory, interestingly enough, is weakest in its own backyard. While market researchers scramble in search of independent variables, they neglect their own dependent variables and usually fail to study the order and organization of consumer tastes themselves. The reason for this may be that market researchers work for a finite set of clients who dictate the dependent variable and have little or no funding, or time for basic research. Most of the recent life style research effort seeks to relate social or psychological traits to one another but fails to relate consumer traits as behaviors to one another in a meaningful causal model.

The most successful modeling of consumer behavior was carried out in a simpler period of history by rural sociologists, who developed scales of socioeconomic status during the 1930's based upon consumer traits of various sorts (Chapin, 1935; Sewell, 1940). The weight of evidence indicates that the American stratification system during the Great Depression was characterized by a fairly clear hierarchy of material life styles strongly reflective of occupational prestige and income. In my own work, I have shown that major social change has probably occurred in American society since that time (Felson, 1974). It appears that the present consumer goods market is probably not as clearly stratified and that stratification probably is only partially manifested in consumer goods. Thus the discontent with so-called social class as an independent variable in market research is partly justified, yet the proper response to that discontent is to build more sophisticated models relating stratification to material life styles, including knowledge, time, family, life cycle and the like. Most important, modern models must consider how one consumer decision affects another. For example, once you buy an 8-room house, you need more furniture, telephones, and consumption of utilities. Once your daughter sets up a separate household in another city

at your expense, you have a whole set of costs tied to that first decision. Once you buy a large car and move 20 miles from work your gasoline consumption is largely determined. Thus, consumer decisions have causal impacts on one another. These impacts can be thought out and then measured. In any case, no one consumer decision should be taken in isolation from others. Clients are perhaps too short-sighted to see this. A large auto manufacturer might be surprised to learn that one of its major competitors is Gerber Baby Food, since families sometimes choose between a new car or a new baby. Yet an auto company is probably too worried about other auto companies as competitors and pay too little attention to the less obvious competition. The latter can be understood only on the basis of understanding the whole consumption process.

I will turn now to a discussion of a current market research fad—life style research. Interestingly, what some of the "life style" adherents see as its greatest strength is actually its Achilles' heel, namely the generation of so-called precise profiles of consumers as individuals. If you really wanted to do the latter, just take interviewer identification code as a set of dummy variables and you can explain 100 percent of the variance in any dependent variable or you can establish that people who read Playboy probably agree with Playboy's ideas. Such information is too specific to generate much understanding, whatever the size of the correlations. Taking 300 independent variables in a sample of 1000, it is little wonder that some researchers have increased their  $R^2$ . Much of this increment is sampling error. The rest is theoretically meaningless. Life style cannot help one to understand consumer behavior if life style variables are a disorderly, non-general, nonhierarchical, or atheoretical set of vaguely related traits whose causal relationships to each other and to anything else are unspecified. Much life-style research could better be termed "idiosyncrasy research," since it uses the computer to group people with similar idiosyncrasies. The central point is that one purchases generality only at the price of specificity and vice versa. Life style research today contributes no understanding with 300 independent variables, unless these can be reduced to a few meaningful dimensions.

Nor does life style research have a right to claim that it is studying "the whole man." No one studies whole men or women. To do so one would have to include psychoanalyses, eye examinations, family tree investigations, and autopsies. The basic research process is one of abstracting certain dimensions from the whole and looking only at those dimensions (i.e., variables). The life style researchers are no different from any other researchers, except that they sometimes abstract 300 variables from the "whole man."

My central point in this paper is to look for a small number of organizing principles which structure consumer behavior, and especially to work out how one type of consumption helps organize another. Fundamental conflict between the short-run goals of business and the long-run goals of science will need to be faced. A good model of family consumption will need to combine and scale products and to study consumption as a system. The prospect of selling more widgets is far from guaranteed. Only basic research can do the job, and it is more likely to come from the ivory tower people than from the Ivory Soap people—though the latter might benefit from it if they wish.

One of the first steps is to abandon discriminant analysis with its multiplicity of dependent variables. Consumer traits will need to be combined into indexes which economize on information. Fancy statistical models

cannot compensate for underdeveloped theoretical models.

To summarize what I have said, modern stratification research has abandoned W. Lloyd Warner for O.D. Duncan; concentrating on multivariate causal modeling, using continuous variables whose effects on one another are decomposed, often with path analysis or other equation systems. Market research needs to take a similar tack and to develop models of material life styles which go beyond the immediate needs of single companies, taking into account at least some of the interrelationships among consumer traits.

#### REFERENCES

- Blau, Peter and Duncan, O.D. The American occupational structure. New York: Wiley, 1967.
- Chapin, F.S. Contemporary American institutions. New York: Harper, 1935.
- Dietrick, B.A. Social mobility: 1969-1973. Annals of the American Academy of Political and Social Science, 1974, 414, 138-147.
- Felson, M. Social Stratification in household attributes: changes since the Great Depression. University of Illinois at Urbana-Champaign, Department of Sociology, Program in Applied Social Statistics. Working Paper #7416, 1974.
- Lasswell, T.E. and Benbrook, S.L. Social stratification: 1969-1973. Annals of the American Academy of Political and Social Science, 1974, 414, 105-137.
- Sewell, W. The construction and standardization of a scale for the measurement of the socio-economic status of Oklahoma farm families. Agricultural and Mechanical College, Stillwater, Oklahoma, Agriculture Experiment Station. Technical Bulletin Number 9, 1940.
- Sheldon, E.B. Family economic behavior: Problems and prospects. New York: Philadelphia; Lippincott, 1973.
- Warner, W.L. Yankee city. New Haven; Yale, 1963.

LIFE CYCLE REVISITED: APPLICATIONS  
IN CONSUMER RESEARCH

Subhash C. Jain<sup>1</sup>  
University of Dayton

This article reviews the Status of Life Cycle as an independent factor in consumer research. Although many studies upheld that life cycle discriminated significantly, its usefulness has been found limited. The logic of using life cycle, which is just one phase of consumer's life style, alone as a predictor over multidimensional approaches, is questioned. A question is also raised over the current practice of constituting life cycle stages, based on age of the head of household, marital status, presence and ages of children.

The concept of family life cycle has been the subject of discussion among marketers for over two decades. A number of studies has found the concept usable in practice for understanding consumer behavior (Carman, 1965; Lansing & Kish, 1957; Lansing & Morgan, 1955; Glick, 1947; Lydall, 1955; Andreassen, 1966). Still other studies have questioned the value of the family life cycle concept in explaining behavior (Rich & Jain, 1968). Lack of any recent, comprehensive, empirical study makes it difficult to arrive at any significant conclusions on the worth of family life cycle in understanding, explaining and influencing consumer behavior. The purpose of this paper is to discuss the conceptual standing of family life cycle, and to critically examine the value of using it as an independent factor in consumer research. The discussion will be substantiated by the findings of two proprietary surveys in which life cycle figured as a predictive variable. The usefulness and acceptance of the concept would also be looked into based on case studies.

#### Definition of Life Cycle

The stage in the life cycle is a refinement of a variable used in analyzing purchasing patterns - age of head of the household. The basic assumption underlying this concept is that for marketing studies, an individual's position in the household life cycle is more important than age or income. For example, the decision of a 40 year old man to buy a house is greatly a function of the fact that he has a wife and three children rather than he is 40 years old. The household life cycle begins when a person leaves the home of his parents and moves into another abode. It proceeds from there to the married stage and finally to the stage of solitary survivor. The married stage can be subdivided into stages marked by the birth of children, specific ages of children, and departure of the children to form their own households. Thus family life cycle may be based on a combination of different demographic variables, such as age of the head of household, marital status, presence and ages of children, and age of the housewife.

Rich and Jain (1968) used age of the head of household and presence of children as the differentiating variables between various life cycle stages. The Life study (National Industrial Conference Board, 1965) formed life cycle

stages based on presences and ages of children. In the survey of consumer finances, the Michigan Survey Research Center, in addition to using age of the head of household and presence of children, included the ages of children in its design (Katona, Lininger & Mueller, 1964).

The question arises which is the most appropriate way to categorize the life cycle. Among the published studies on the subject, no two follow the same classification. This makes comparison difficult. But age of the head of household achieves significance since age as a natural process would make a difference in behavioral patterns irrespective of the family cycle status (Fisk, p.118). For example, a married individual will be a different person in early 20's than when he is over 40, even if no children are involved. Granbois (1963) found that degree of interaction in the process of household decision making between husband and wife declines as the couple becomes older. Webster (1965) discovered the older housewives to be more deal-prone. They bought fewer units and many more brands. Ages of either spouse, however, need not be included in constituting life cycle stages if the ages at which people get married and children are born are significantly consistent among the majority.

Marriage obviously would change one's life perspective. Myers and Reynolds (1966) mentioned the results of a proprietary study that showed how marital status became an intervening variable in the preference for a dressed-up engine with chromed valve covers of an automobile with the married men preferring less glamorous options. Likewise, presence of children would have an effect on a family's life style. One study noted that along with factors such as perceived consumer effectiveness, the presence of children in the family increased the probability of using returnable bottles and nonphosphate laundry detergent (Kinneary, Taylor & Ahmed, 1972). Another study indicated that married couples with children attached less importance to having a home closer to the family head's place of work than singles and couples without children (The Outlook on Consumer Behavior, 1964). Katona (1960) mentioned that young couples with children engaged in installment buying more frequently. Montgomery (1970), however, discovered that the presence of children had little effect on dealing activity by the family.

Another variable that has found place in forming life cycle groupings is the ages of children. Disagreement prevails among researchers whether the age of the youngest child be considered, or that of the oldest, or those of both (Frank, Massy & Wind, 1972, p. 34). Intuitively, inclusion of age of children in the analysis should make life cycle categories too small to be meaningful from the viewpoint of a practical manager. As a matter of fact, the number of children also may be considered in forming life cycle stages. Many studies have found family size as a viable factor in explaining consumer behavior. For example, Crockett & Friend (1960) found family size to be one of the most highly correlated variables with overall food consumption. A similar conclusion was drawn by Frank, Massy & Boyd (1967). Here again the danger is if along with ages of children, the number of children is used for constituting life cycle stages, each stage may still become smaller.

The definition of life cycle as discussed above attaches little significance to income or any other economic variable. Thus it is assumed that married families with younger children would behave alike, even if they are located on different steps on the status ladder. This assumption hardly appears reasonable. Engel, Kollat & Blackwell (1968) indicated that the influence of social class on consumer behavior varied in different stages of the life cycle. Accepting their hypothesis, to make life cycle stages meaningful, each stage may further be split into status categories. Once again we may be faced with the problem that a status-life cycle category may become too narrow to be of value. As has been said: "If a category

is too narrow, it will include such a small proportion of the sample that it will be all but unpopulated except in large surveys " (Wells & Gubar, 1966, p. 360).

In the final analysis, the bases used to break down family life cycle will have to depend on two factors: nature of the product, and prerequisites of a good segmentation criterion. In the case of products such as baby foods, school supplies, vacation, entertainment, toys, and large appliances, perhaps inclusion of ages of children as a criterion for forming life cycle stages would be desirable. For segmenting the market for products such as housing, furniture and furnishings, insurance, small appliances, and cars, the variables marital status, age of the head of household, and presence of children may suffice. Of course, different criteria may have to be used for individual products within the same broad group. To segment the market for appliances, as an example, generally speaking ages of children would be important. But ages of children may be more meaningful for discriminating buying of refrigerator, vacuum cleaner and washing machine than air conditioner, electric floor waxer, and compactor, since purchase of these items may depend more on the level of discretionary income than the ages of children. As far as the criteria for good market segmentation are concerned, the following conditions should be met:

- (a) Each segment should show varied demand elasticities in response to price and promotional variables (Engle, Fiorillo & Cayley, 1972; Frank, 1968).
- (b) Each segment should differ in the average purchase rate of the product being considered.
- (c) Within a segment the heterogeneties should be minimal.
- (d) Each segment should be large enough to justify a separate marketing program (Bieda & Kassarian, 1969; Engel, Kollat & Blackwell, 1969).
- (e) Each segment should be within the reach of the firm's promotional program.

#### Applications

Many consumer researches have been reported in literature in which life cycle figured as an independent variable. A comprehensive summary of these studies has been prepared by Wells & Gubar (1966). However, a brief review would be desirable here. Applying the concept of family life cycle to consumer finances, Lansing & Morgan (1955) discovered that people differed in their behavior patterns in different stages of family life cycle with respect to income, expenditures on durables, assets and debts, and subjective feelings about financial position. For example, the income of a typical family increased until the head of the household reached late forties, and declined thereafter, with the exception of the young, married, and children under six stage, since during this stage the wife stopped working.

Social Research, Inc. (1964) in a study undertaken for the Kroehler Manufacturing Company, noted that interest in furniture buying was highest among newly married couples and among families with grown-up children. The young family, however, placed relatively greater emphasis on sensibility and practicality in furniture buying, while the families in the later stages attached more importance

to attractiveness and reflection of good taste. Barton (1955) found that life cycle discriminated in behavior better than age in the case of nondurable goods. For example, young housewives tended to buy prominently advertised products, products in smaller sizes and packages and new products. Older housewives, however, bought larger sizes and multiple packages, were less likely to be influenced by advertising and showed less interest in new products. A more recent evidence of the link between a family's stage in the life cycle and consumer traits was found in the National Industrial Conference Board (NICB) study (1965), sponsored by Life, on the Expenditure Patterns of the American Family. This study provides a most comprehensive treatise on life cycle, covering expenditures on services, and nondurable and durable goods. The NIBC Study showed that expenditures varied by age and life cycle to about the same degree for about 50 percent of products and services. But for 177 products and services life cycle was found to be more discriminating than age which excelled in 54 cases.

One study has taken an exception to the findings of the studies reported above. Rich & Jain (1968) in their study of shopping behavior questioned the applicability of life cycle as a predictor of shopping behavior. The authors, based on the data of an extensive empirical study of women's shopping behavior, suggested that life cycle distinctions were obscured by the leveling effects of socioeconomic changes. Similar comments were made by a baking company executive on the usefulness of life cycle:

Basically, from a marketing standpoint, bread is considered a staple - it's on the housewife's list, and the question is, which type of bread and brand she will select. Snack cakes are primarily an impulse item, and the decision to buy is made in the store more often than not.

As far as bread is concerned, it is generally agreed that children and young people prefer soft white bread. As people grow older, they tend to prefer a firmer bread. However, there appear to be geographic differences. In the South and Midwest, soft bread is asked for. On the eastern seaboard and the Far West, more of the firmer bread is sold. Also, the East and Far West consume more variety breads - rye, whole wheat, French, Italian and so on, than in other sections.

To try to break consumption down into the categories you mention would be most difficult. Obviously, the larger the family, the greater the consumption. But it might be a mistake to assume that per person consumption for your group 2 is markedly different from that of, say, group 4, since both parties in the newly married couple will tend to eat away from home more often, and their restaurant or sandwich shop consumption includes bread which otherwise would be purchased for the home. Another difficulty is that national eating habits are changing. We are snacking more and eating fewer full meals. Also, far more meals are eaten away from home than ever before - not only in restaurants and coffee shops, but in schools, day care centers, industrial feeding establishments, homes for the elderly, hospitals, and so on. As for snack cake, consumption probably rises rapidly from kindergarten, levels off in the mid-teens, declines in the late teens and twenties, though it by no means disappears.

The studies referred above have been based on information gathered in 1950's and 1960's. In the absence of any recent empirical work, verification of earlier studies on their standing today must be made on indirect evidence.

Table 1 shows the results of a survey on shopping behavior, where both life cycle and age were used as independent variables. In most cases, it would be noted, the chi-square values for age are considerably larger than those for life cycle. In quite a few cases, age chi-square values were far in excess of those for life cycle, for example, persons with whom shopping discussion was found helpful. The life cycle was favored for nature and extent of discount shopping, services expected from a department store, extent of shopping done by mail and fashion consciousness. On balance, it might be reasonable to conclude that neither age or life cycle is able to predict shopping behavior to any large extent. Further use of age in segmentation may be as useful as life cycle.

A similar conclusion would be drawn from the findings presented in Table 2. Usage of 22 of the 36 financial services tested was found to vary among respondents when categorized by sex, age, educational level, income level, and family life cycle. Of these, usage varied by stage of family life cycle alone for 14 services. This was followed by age (13 services) and by educational level and income level (both with 10 services).

### Conclusion

Life cycle, it should be recognized, is but one of the many variables that affect consumer behavior. The entire spectrum of behavior cannot be explained by a single variable. It is the total life style that would determine how a group of people may behave in the market place. According to Plummer (1974), activities, interests, opinions, and demographics are the four major lifecycle dimensions, with 9 variables under each dimension, making a total of 36 variables. Life cycle falls under demographics along with 8 others. Clearly, life cycle alone cannot be expected to discriminate in behavior, assuming all other life style variables to be constant. There may be situations where life cycle, as also income, social class, or age, may reveal significant differences in consumer behavior. But such differences may lack reliability.

Further, it would be easy to make a case for life cycle through discovering broad distinctions among consumers in different stages of family life cycle. But how far such distinctions are usable at the micro level is difficult to say. Consider Table 3, as an example. Based on information collected from six hospitals, the health care needs in different stages in family life cycle have been compiled. Obvious differences are noticeable in various life cycle stages. Health care requirements are minimum in the bachelor stage. In the later stages, the families tend to require more intensive medical care with the exception of older couples, with grown up dependent children. It is questionable, however, if this type of information might help a drug company in formulating its marketing strategy, or a hospital in programming its service offerings.

As an hypothesis, it appears that the concept of family life cycle, as has been conceived and used so far, may be outdated. The concept, however, may be made more viable through using a combined life cycle-socioeconomic index, built on life cycle, income, social class, family size, etc. Also attempts might be made for inclusion of households in the analysis which past studies on life cycle have ignored. These are middle aged single people, either never married or divorced or widowed; and older married couples who never had children. Possibly one way of improving the present categorization of family life cycle would be to use three different life cycles: the bachelor (bachelorette) life cycle; the divorced/widowed life cycle;



TABLE 1

Chi-Square Values for Life Cycle and Age Groups

Shopping Behavior Aspects	Life Cycle $\chi^2$	Age $\chi^2$
Shopping attributes one likes most	8.32	23.24 <sup>b</sup>
Shopping attributes one likes least	19.01	36.81 <sup>b</sup>
Extent to which shopping is liked	16.47	41.16 <sup>b</sup>
Frequency of shopping	21.52 <sup>b</sup>	6.15
Extent of downtown shopping currently	4.44	
Extent of downtown shopping in the past	9.29	
Items bought in discount stores	948.87 <sup>a</sup>	252.61 <sup>a</sup>
Extent of discount store shopping	42.45 <sup>b</sup>	8.41
Type of store where easier to shop	7.12	
Services considered important in a regular department store	25.15 <sup>b</sup>	2.63
Extent of shopping done by telephone	18.51	17.89
Extent of shopping done by mail	67.70 <sup>a</sup>	24.32 <sup>b</sup>
Frequency with which newspaper ads are looked at	8.68	8.04
Newspapers in which ads are looked at	3.11	11.47
Helpfulness attributed to newspaper ads	16.53	13.88
Fashion consciousness	210.05 <sup>a</sup>	119.46 <sup>a</sup>
Methods helpful for keeping track of fashion trends	752.69 <sup>a</sup>	938.10 <sup>a</sup>
Extent to which shopping discussion with others is helpful	13.08	8.61
Persons with whom shopping discussion is helpful	41.19 <sup>b</sup>	125.08 <sup>a</sup>
Extent to which shopping is done with others	5.12	7.59
Shopping companions	158.53 <sup>a</sup>	288.17 <sup>a</sup>
Comparison shopping	1.42	5.73
Browsing through stores	6.46	6.94
Bargain hunting	10.60	9.39
Impulse buying	11.34	13.42
Combining shopping with grocery buying	3.66	7.88
Extent of decision making about what to buy	16.63	12.41
Day and time preferences for shopping	8.16	38.82 <sup>b</sup>
Number of charge accounts in stores	157.38 <sup>a</sup>	18.86

<sup>a</sup> significant at .01 level or beyond

<sup>b</sup> significant at .05 level or beyond

Note: Above results are based on 840 interviews with women shoppers using quota sampling. The size and number of the life cycle and age segments was equal.

Source: Proprietary study sponsored by a major department store.

TABLE 2

Life Cycle, Compared to Other Socio-Economic Variables, as a  
Criterion for Market Segmentation: Use of Banking Services

Market Segmentation Variable	Sex	Age	Education Level	Income Level	Stage of Family Life Cycle
Financial Service					
Regular checking account				YES	YES
Savings account		YES	YES		YES
Cashing paychecks		YES	YES	YES	
Cashing personal checks			YES		
Drive-In banking		YES		YES	YES
Safe deposit box		YES		YES	YES
Bank by mail		YES	YES	YES	YES
Bank credit card		YES	YES	YES	YES
Travelers checks			YES	YES	YES
Auto loan		YES			YES
Personal loan		YES			YES
Home mortgage loan		YES	YES		YES
Christmas club					
Personal credit reference		YES			YES
Automatic payroll deduction	YES	YES			YES
Bill payment		YES			YES
Certified checks					
Money orders					
Time certificates of deposit					
Special checking accounts					
Purchase stocks and bonds			YES		
Investment advisory service					
Home improvement loan		YES			
Farm or business loan				YES	
Education loan					YES
Purchase mutual funds			YES	YES	
Insurance on a loan				YES	
Checking account with credit reserve					
International banking					
Corporate service					
Travel or vacation loan					
Trust services			YES		
Data processing					
Pension and profit sharing plan					
Boat loan					
Mobile home loan					

Note: "YES" indicates that a particular variable significantly segments the market with reference to a particular financial service.

Source: King, A. L., McMillan, J. R., & Merenski, J. P. Market Segmentation Based on Usage of Banking Services. University of Dayton, School of Business Administration. Unpublished paper, 1974.

TABLE 3

Family Life Cycle And Health Care Services Needs/Purchasing Habits

Stage In Family Life Cycle

Bachelor Stage- (young single people)	Newly Married Couples - (young, no children)	Full Nest I- (young married couples with youngest child under six)	Full Nest II (young married couples with youngest child six or over)	Full Nest III- (older married couples with dependent children)	Empty Nest - (older married couples with no children living with them)	Solitary Survivors - (older single people)
<p>Episode health care. No personal physician. Small percentage of income used to purchase health care services or health insurance coverage. Need for services not significant - good health maintained without many services. Little concern about health or health coverage.</p>	<p>More regular care. Find family physician and specialty physician (i.e., Obstetrician, Gynecologist). Greater percentage of income spent for services and coverage than during bachelor stage. Still no real need for acute services such as hospitalization. Good health maintained easily. Some dental services used. Still unconcerned about health, and health coverage has low priority.</p>	<p>More health services being purchased from more providers: hospital, Physicians; family M.D.; Obstetrician; Gynecologist; Pediatrician; possibly a Podiatrist for children's shoes. Greater percentage of family income devoted to purchasing services and coverage. Health of family still easily maintained.</p>	<p>"Middle class medicine" important-i.e., a specialist for each set of problems: Obstetrician, Gynecologist, Pediatrician, Ophthalmologist, Orthodontist. Medical emergencies more common - therefore, more use of acute services. Greater percentage of income for services and coverage. Family's health still relatively easy (inexpensive) to maintain.</p>	<p>Health maintenance primary concern. Annual physicals and episodic care as needed. Major health care expenses associated with childhood and medical emergencies are infrequent. Constant or declining percentage of income (compared with Full Nest II) needed for health care services and coverage.</p>	<p>Health maintenance primary concern. Some additional purchasing of services for corrective lenses, dental plates, and hearing aids. Increasing need for health services for chronic conditions (example: arthritis, ulcers, etc.) Health not maintained as easily without services.</p>	<p>Health care service needs are varied: acute care-hospital; chronic care-physician, use of custodial or convalescent care-nursing home. Views visits to physician as "something to talk about." More real or imagined need for services. Large percentage of income needed for health care services.</p>

and the married life cycle. Within each stage, there will be a group of situations that would affect consumer behavior. Thus each cycle may be broken into various stages based on a relevant "situations vector."

The literature search shows that linking of market behavior to family-age and child-rearing, i.e., family life cycle, was useful in the realm of housing purchases, finances, and the purchase of certain packaged consumer goods (Hisrich & Peters, 1974). Other evidence suggests, however, that life cycle may have lost significance as a predictive variable. A variable significant in one market/product context may not be significant in another. Possibly a predictive variable needs to be tested in each individual market/product context to ascertain its usefulness.

Years ago Gary Steiner (1966, p. 211) mentioned a few propositions on consumer behavior. These propositions are equally true today. We may, therefore, conclude our discussion on life cycle with these propositions: (1) "Some do, some don't," (2) "The differences are not very large," and (3) "It is not as simple as that."

#### FOOTNOTE

<sup>1</sup>Subhash C. Jain is Associate Professor of Business Administration, University of Dayton.

#### REFERENCES

- Andreasen, A. R. Geographic mobility and market segmentation. Journal of Marketing Research, 1966, 3, 341-348.
- Barton, S. G. The life cycle and buying patterns. In L. H. Clark (Ed.), Consumer behavior. Vol. 2. New York: New York University Press, 1955.
- Bieda, J. C. & Kassarjian, H. H. An overview of market segmentation. In Barnard A. Morin (Ed.), Marketing in a changing world. Chicago: American Marketing Association, 1969, 249-253.
- Carman, J. M. Studies in the demand for consumer household equipment. Berkeley: Research Program in Marketing, Institute of Business and Economic Research, University of California, 1965.
- Crockett, J. & Friend, I. A Complete set of consumer demand relationships. In I. Friend & R. Jones (Eds.), Consumption and saving. Vol. I. Study of consumer expenditures, incomes and savings. Philadelphia: University of Pennsylvania Press, 1966.
- Engel, J. F., Fiorillo, H. F., & Cayley, M. A. (Eds.) Market segmentation: Concepts and applications. New York: Holt, Rinehart, & Winston, 1972.
- Engel, J. F., Kollat, D. T., & Blackwell, R. D. Consumer behavior. New York: Holt, Rinehart, & Winston, 1968.
- Engel, J. F., Kollat, D. T., & Blackwell, R. D. Personality measures and market segmentation. Business Horizons, June, 1969, 12, 61-70.

- Fisk, G. Marketing Systems: An introductory analysis. New York: Harper & Row, 1967.
- Frank, R. E. Market segmentation research: Findings and implications. In F. M. Bass, C. W. King, & E. A. Pessemier (Eds.), Applications of the sciences in marketing management. New York: John Wiley, 1968, 39-68.
- Frank, R. E., Massy, W. F., & Boyd, H. W. Correlates of grocery product consumption rates. Journal of Marketing Research, 1967, 4, 184-190.
- Frank, R. E., Massy, W. F., & Wind, Y. Market segmentation. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1972.
- Granbois, D. H. The role of communication in the family decision-making Process. In S. A. Greyser (Ed.), Toward scientific marketing. Chicago: American Marketing Association, 1963, 44-57.
- Hisrich, R. D. & Peters, M. P. Selecting the superior segmentation correlate. Journal of Marketing, July, 1974, 38, 60-63.
- Katona, G. The powerful consumer. New York: McGraw-Hill Book Company, 1960.
- Katona, G., Lininger, C. A., & Mueller, E. 1963 Survey of consumer finances. Michigan Survey Research Center, University of Michigan, 1964, Monograph No. 34.
- Kinnear, T. C., Taylor, J. R., & Ahmed, S. A. Socioeconomic and personality characteristics as they relate to ecologically constructive purchasing behavior. In M. Venkatesan (Ed.), Proceedings. Third Annual Conference, Association for Consumer Research, 1972, 34-60.
- Lansing, J. B. & Morgan, J. N. Consumer finance over the life cycle. In L. H. Clark (Ed.), Consumer behavior. Vol. 2. New York: New York University Press, 1955.
- Lansing, J. B. & Kish L. Family life cycle as an independent variable. American Sociological Review, 1957, 22, 512-519.
- Lydall, H. F. The life cycle in income, savings, and asset ownership. Econometrica, 1947, 12, 164-174.
- Montgomery, D. B. Consumer characteristics and 'deal' purchasing: An example from the dentifrice market. Marketing Science Institute. Working paper, 1970.
- Myers, J. H. & Reynolds, W. H. Consumer behavior and marketing management. Boston: Houghton Mifflin Company, 1966.
- National Industrial Conference Board. Expenditure patterns of the American family. New York: Life, 1965.
- The outlook on consumer behavior. Ann Arbor, Michigan: The Foundation for Research on Human Behavior, 1964.

- Plummer, J. T. The concept and application of life style. Journal of Marketing, January, 1974, 38, 33-37.
- Rich, S. U. & Jain, S. C. Social class and life cycle as predictors of shopping behavior. Journal of Marketing Research, 1968, 5, 41-49.
- Social Research, Inc. Furniture buying and life stages. In M. H. Grossack (Ed.), Understanding consumer behavior. Boston: The Christopher Publishing House, 1964, 287-290.
- Steiner, G. A. Consumer behavior: Where do we stand? -- A psychologist's appraisal. In J. W. Newman (Ed.), On knowing the consumer. New York: John Wiley, 1966.
- Webster, F. E., Jr. The 'deal-prone' consumer. Journal of Marketing Research, 1965, 2, 186-189.
- Wells, W. D. & Gubar, G. Life cycle concept in marketing research. Journal of Marketing Research, 1966, 3, 355-363.



THE STATUS OF CONSUMER BEHAVIOR:  
SOME EMPIRICAL PERSPECTIVES

Gary T. Ford and Philip G. Kuehl  
University of Maryland  
Robert F. Dyer<sup>1</sup>  
The George Washington University

The state-of-the-art dimension of consumer behavior has received increased discussion in the consumer behavior literature in recent years. This paper provides empirical insights into three prominent issues in consumer behavior obtained from a sample of Association for Consumer Research members. The three issues investigated in this study include: (1) the interdisciplinary heritage of consumer behavior; (2) the development and use of consumer behavior research and (3) present needs of and future directions for consumer behavior research.

Introduction

The state-of-the-art of consumer behavior has been discussed by several authors in recent years. However, a review of these papers suggests that divergent opinions exist among consumer behaviorists regarding the development of the concepts, theory and practices which encompass the present status of the discipline. This paper reports some empirical perspectives on three important issues which are frequently discussed in state-of-the-art reviews: (1) the interdisciplinary heritage of consumer behavior; (2) the development and use of consumer behavior research and (3) present needs of and future directions for consumer behavior research. The findings summarized in this paper were obtained through a survey of the membership of the Association for Consumer Research conducted in November, 1973. Since ACR members have a professional interest in the field, it is believed that a survey of these individuals is an acceptable approach for examining the status of consumer behavior by identifying opinions on the issues addressed in this study.

The findings contained in this study provide potentially useful insights for three groups of consumer behavior professionals. For corporate, public policy making, and consumer advocate users of consumer behavior research, findings discussed here highlight the theoretical assumptions and practical limitations inherent in applying this body of knowledge to decision-making environments. For academicians and practitioners interested in advancing consumer behavior theory, perspectives generated in this study contain insights which can be integrated with future conceptual developments in the field. Finally, this study assists consumer behavior practitioners, theorists and researchers to define three aspects of the existing status of the field from an empirical perspective. Such a viewpoint is particularly appropriate for consumer behavior--a subject of study that has experienced a two fold increase in its literature base during the last five-year period [Kollat, Blackwell and Engel, 1972].



This paper is organized in the following manner. First, the theoretical framework encompassing the survey is presented and research questions which define the scope and thrust of the study are identified. Second, the research methodology and characteristics of the sample are described. Third, the results and findings from the survey are presented and, finally conclusions and implications of the study are discussed.

#### Theoretical Framework

An examination of state-of-the-art reviews by Tucker [1967], Sheth [1972], Engel, Kollat and Blackwell [1972, 1973] and Ward and Robertson [1972] stimulated the identification of the three major issues examined in this paper. A summary of the positions taken in each of these earlier papers on these three issues are presented with accompanying research questions in this section.

#### 1. The Interdisciplinary Heritage of Consumer Behavior

It is generally recognized that intensive interdisciplinary borrowing from the social sciences has been characteristic of and contributed immeasurably to the development of the body of knowledge of consumer behavior. This tradition of borrowing has in a very real sense created an identity crisis in consumer behavior and has stimulated discussion in two areas regarding: (1) where consumer behavior fits in relation to marketing and the behavioral sciences and (2) whether consumer behavior will return more to the other disciplines than it now borrows.

Tucker [1967] suggests that consumer behavior is a necessary dimension of the marketing process and, yet, is not independent of the behavioral sciences. Ward and Robertson [1973] believe that consumer behaviorists have borrowed almost all constructs from the behavioral sciences while attempting to understand behavior in the consumption role. Sheth, [1972], on the other hand, believes that consumer behavior is neither a subset of marketing nor a subset of the social sciences but is its own unique discipline. These viewpoints suggest that agreement does not exist among consumer behaviorists of the normative identity of the field or its relationship to marketing and the established social sciences.

Another issue of concern to consumer behaviorists is whether other disciplines will begin to utilize concepts and techniques originated in consumer behavior. Sheth [1972] believes that borrowing from consumer behavior will occur in the less mature social sciences, in the hard sciences and especially in the social sciences which he says are facing crises of relevance. Tucker [1967] feels that consumer research and behavioral science research are coextensive and cosupportive. Ward and Robertson [1973] state that ideally consumer behavior research should add to the cumulative knowledge of the fields that provided the conceptual framework for the study, but appear pessimistic when assessing whether consumer behavior has contributed much to other fields.

The positions taken by these authors on the two major issues concerned with the interdisciplinary heritage of consumer behavior led us to investigate the following questions:

- (1) Is consumer behavior more of a behavioral science discipline than a subset of marketing?
- (2) Will consumer behavior eventually contribute more to the social sciences than it now borrows?

- (3) Has the study of consumer behavior really strengthened the body of knowledge of marketing?

## 2. The Development and Use of Consumer Behavior Research

A major issue which has permeated the development of consumer behavior concerns research practices in the discipline. State-of-the-art papers have generally castigated consumer behaviorists for their research practices ranging from borrowing irrelevant concepts from the behavioral sciences to improper application of consumer behavior research findings. For example Ward and Robertson [1973] believe (1) that consumer behavior research is replete with unsystematic and irrelevant borrowing from the behavioral sciences and (2) that often times the assumptions of theories and constructs borrowed from the social sciences are not totally satisfied. Tucker [1967] comments that borrowing has been incomplete and very specific and that consumer behavior research results are of limited utility because of this. Engel, Kollat and Blackwell [1973] state that when theories are used properly they have led to insightful results but that too often theories are used improperly. Ward and Robertson [1973] see a danger in researchers concentrating too much on immediate applications for business without emphasis on the development of understanding. Tucker [1967] exhibits similar concerns when arguing that consumer behavior researchers should not confuse marketing theory with corporate practice.

These statements exhibit a common concern among state-of-the-art reviews regarding the development of consumer behavior research practices and led us to investigate the following questions:

- (1) Do consumer behavior researchers borrow from other disciplines without properly evaluating assumptions.
- (2) Do consumer behavior researchers rely on irrelevant or inappropriate concepts for the questions being investigated?
- (3) Are consumer behavior models used properly or improperly by consumer behavior researchers?
- (4) Is there a tendency for businessmen to misuse the results of consumer research?
- (5) Would research findings have been different had they been based on comprehensive models?

## 3. Present Needs of and Future Directions for Consumer Behavior

When identifying the present needs of consumer behavior and future directions in which the discipline is likely to move, state-of-the-art papers discuss the following two issues: (1) long run theory development versus immediate applications, and (2) the need for and prospects for the development and evaluation of comprehensive models of consumer behavior.

Ward and Robertson [1973] have stated that a basic conflict exists among the users and producers of consumer behavior research regarding whether more time should be spent on understanding relationships or on applications. Tucker [1967] recognizes the unique relationships and obligations existing between consumer behaviorists and businessmen but believes more time must be spent on developing more complete models.

Engel, Kollat and Blackwell [1973] concur with Ward and Robertson [1973] who state that consumer behavior research faces a "pay off dilemma" in that it has not led to any comprehensive and validated theories and simultaneously has not been very helpful in applications for business.

What then are the current needs of and prospects for the development of consumer behavior research? Ward and Robertson [1973] believe that the greatest hope for progress in the consumer behavior field lies in the future development of middlerange theories. Kollat, Blackwell and Engel [1972], however, believe that consumer behaviorists must begin relying on comprehensive models. Furthermore, they would not be surprised if over 90 per cent of research results would be different had they been based on complete models of consumer behavior. Therefore, Kollat, Blackwell and Engel see as a present need the development of comprehensive models of consumer behavior, while Ward and Robertson feel that the development and use of comprehensive models is premature. Sheth [1972] concurs with Kollat *et. al.* and sees a re-emergence of quantitative model building in consumer behavior research. Kollat *et. al.* also comment on the need for statistical evaluation of comprehensive models. Tucker [1967] believes that measurement problems must be solved before comprehensive models can be developed, while Sheth [1972] predicts that the measurement problems will be reconciled in five years. Kollat, Blackwell and Engel [1972] and Ward and Robertson [1973] appear to be less optimistic regarding the resolution of measurement problems than is Sheth.

It is clear from this discussion that the problems of (1) the development of consumer behavior theory versus the applications desires of research users and (2) the divergent viewpoints regarding the best way to advance the discipline of consumer behavior have been of utmost concern in state-of-the-art papers. With these points in mind the following research questions were investigated:

- (1) Should consumer behavior researchers be more interested in identifying and understanding relationships or in applications for business?
- (2) Should researchers be working on comprehensive models or on middle-range theories?
- (3) Must measurement problems be reconciled before comprehensive models can be developed?
- (4) What are prospects for the development of criteria and standardized tests to evaluate comprehensive models of consumer behavior.

The preceding discussion has highlighted the theoretical framework of and the type of research questions investigated in this study. Findings presented in the subsequent sections provide an indication of the opinions of a sample of ACR members regarding these issues.

#### Methodology

To evaluate the status of the issues investigated in this study, a structured mail questionnaire was sent to 463 ACR members. A total of 277 questionnaires (59.8 percent) were returned and 258 (55.8 percent) of these were usable for data analysis. Since this was intended to be an exploratory empirical study of the status of consumer behavior, this response rate was regarded as satisfactory and no follow-up procedures were employed.

The sample respondents included professionals from a variety of academic disciplines concerned with consumer behavior research and theory. For instance, 55.4 percent of the respondents held degrees in marketing, 12.0 had training in psychology, 7.3 percent were behavioral scientists, and 6.1 percent held economic degrees. Fifty-five point nine percent of the sample held doctorate degrees, and an additional 21.4 percent of the respondents had earned masters degrees. Occupationally, 60.9 percent of the respondents are academics, 16.6 percent are business executives, 13.0 percent are graduate students, and 5.1 percent are independent business consultants.

#### Findings

The percentage of respondents agreeing or disagreeing with Likert-type statements concerned with the interdisciplinary heritage and development of consumer behavior is summarized in Table 1. Factor analysis, as described in the Appendix, was used to construct and verify the consistency of the

TABLE 1

#### The Heritage and Development of Consumer Behavior

	Total Number of Respondents	Agree or Strongly Agree	Disagree or Strongly Disagree	No Response or No Opinion
1-1. Consumer behavior is more behavioral science than marketing	100% (254)	53.1%	41.0%	5.9%
1-2. Over the next three years the behavioral sciences will continue to be the dominating influence on consumer behavior	100% (255)	71.4%	21.2%	7.4%
1-3. Consumer behavior should be given credit for showing applications of concepts and theories of other disciplines	100% (256)	69.9%	20.7%	9.4%
1-4. Consumer behavior will eventually contribute more to the social sciences than they now contribute to consumer behavior	100% (256)	19.5%	57.5%	23.0%
1-5. The study of consumer behavior has immeasurably strengthened the foundations of marketing	100% (254)	63.0%	30.7%	6.3%

Likert statements which are grouped for discussion purposes in each of the three tables contained in this article. Table 1 results give an indication of

the opinions of a sample of ACR members regarding where consumer behavior fits relative to marketing and the social sciences and its contributions to marketing and the social sciences. Table 1 shows that a majority of the respondents view consumer behavior as more closely aligned with the behavioral sciences than with marketing. Similarly a majority of the respondents also believed that consumer behavior will also borrow more from the behavioral sciences than it will contribute. However, it was also felt that consumer behavior has contributed both to marketing and to the behavioral sciences.

The statements contained in Table 1 were further analyzed according to those educated in marketing and those educated in other fields. These results indicated that 42.1 percent of the marketers and 72.7 percent of the other field majors agreed that consumer behavior is more behavioral science than marketing. Also a greater percentage of marketers (79.8) than other field majors (66.0) disagreed that consumer behavior will contribute more to the behavioral sciences than it now borrows. Therefore, it appears that consumer behaviorists with a marketing background view their discipline as a marketing discipline which will continue to borrow concepts from the other social sciences for application to marketing problems. Furthermore, answers to where consumer behavior fits relative to marketing and the social sciences appear linked to the respondents training.

Table 2 reports findings regarding ACR members opinions on consumer behavior research practices. The responses to the statements contained in Table 2 indicate that consumer behaviorists generally agree that problems exist with consumer behavior research practices. Furthermore, since the statements investigate different stages of the research process from improper and irrelevant borrowing of concepts and constructs of the social sciences through the improper use of consumer behavior models and to potential incorrect applications by business, it is fairly clear that the respondents believe the research practices problems are extensive. In addition, the results show that consumer behaviorists believe that research based on comprehensive models would have led to different results than those evolving through the application of partial or mid-range models to research processes. Also, the statements from Table 2 were analyzed by years studying consumer behavior, educational backgrounds and occupational areas of the respondents and no substantial differences were detected among respondent groups. In fact 56.0 percent of the businessmen agree that businessmen might misuse research by taking it out of its original context.

The statements contained in Table 3 provide information about ACR members opinions on present needs and future directions of consumer behavior research. The findings indicate that consumer behaviorists should simultaneously spend more time on understanding relationships and be attentive to the practical uses of their studies. Apparently the "pay off dilemma" mentioned by Ward and Robertson [1973] is well recognized by consumer behaviorists. The responses also show that consumer behaviorists should not attempt to start predicting the future behavior of consumers until greater theoretical and conceptual advances are made.

If a general belief exists that comprehensive models would have led to different findings, what are the prospects for some of the major problems impeding the development of comprehensive models? Generally it is believed the more complex models will be developed in the future and that approaches to measurement and evaluation problems will be developed. Furthermore when the responses were analyzed by years studying consumer behavior, educational

TABLE 2

## The Development and Use of Consumer Behavior Research

Statement	Total Number of Respondents	Agree or Strongly Agree	Disagree or Strongly Disagree	No Response or No Opinion	
2-1. There is a danger that consumer behaviorists borrow from other disciplines without properly evaluating inherent assumptions	100.0% (255)	84.0%	10.2%	5.8%	
2-2. Too many times researchers fail to properly examine the relevance of behavioral constructs being used in their consumer research	100.0% (255)	81.2%	7.5%	11.3%	
2-3. There are too many instances in which consumer behavior constructs and theories have not been used properly in research	100.0% (257)	72.7%	12.9%	14.4%	
2-4. If past consumer research would have been based on a comprehensive model, many significant and non-significant findings would have changed	100% (256)	61.3%	17.2%	21.5%	
2-5. A major problem in consumer behavior research is businessmen taking the research out of its original context and thereby misusing the results	100% (254)	63.0%	32.3%	4.7%	
2-6. Most present models of the consumer behavior process are:	<u>Totals</u> 100.0%	<u>Used Properly</u> 17.5%	<u>Neutral</u> 21.1%	<u>Not used properly</u> 47.4%	<u>No opinion or no response</u> 14.0%

backgrounds and occupation of the ACR members the only large difference in responses was with statement 3-1, where 44.7 percent of those in business and 65.1 percent of the academics felt more time should be spent on understanding relationships. Therefore, the percentage responses shown in Table 3 hold across respondents with different amounts of experience, educational backgrounds and occupations.

TABLE 3

## Present Needs and Future Directions of Consumer Behavior Research

	Total Number of Respondents	Agree or Strongly Agree	Disagree or Strongly Disagree	No Response or No Opinion
3-1. Consumer behavior researchers should spend more time on understanding relationships rather than on immediate applications for business firms (256)	100%	51.2%	33.2%	15.6%
3-2. Researchers in consumer behavior must remember to pay heed to the practical marketing applications of their studies (254)	100%	68.1%	26.8%	5.1%
3-3. It is now time for consumer behavior researchers to concentrate on predicting future behavior rather than continuing explaining past or present behavior (256)	100%	48.8%	39.1%	12.2%
3-4. Measurement problems must be reconciled before a comprehensive theory of consumer behavior can be developed (256)	100%	68.4%	24.2%	7.4%
3-5. Over the next five years criteria will be developed to evaluate theories of consumer behavior (258)	100%	57.0%	27.9%	15.1%
3-6. Over the next five years, standardized tests to measure consumer behavior constructs will be developed (258)	100%	53.9%	32.2%	13.9%
3-7. Over the next five years, complex inductive models of consumer behavior will be constructed with the use of statistical procedures (258)	100%	58.5%	26.7%	14.8%

## Conclusions and Implications

There are several important conclusions that can be drawn from this empirical evaluation of the status of consumer behavior. One of the most important findings is the relatively high degree of consensus found among ACR members concerning the three major issues explored in this study. It may be argued that ACR members are somewhat homogeneous and it is not surprising

that they agree on these points. However, we believe that one of the marks of a maturing discipline is common recognition of the current status of the field, of the problems which need to be resolved and of the prospects for their resolution. Remarkably, the sample respondents did exhibit consensus on almost every point being investigated. Furthermore, when the Likert statements were investigated for likely differences in responses due to occupation, educational background or years studying consumer behavior in most cases, no significant differences (chi squared) were found. This result implies that the opinions exhibited in the tables are not substantially different from different individuals but are recognized in common by the respondents to this study.

Sample respondents generally held similar beliefs on issues related in the interdisciplinary heritage of consumer behavior. While interdisciplinary borrowings have been utilized in an applied context by consumer behavior specialists, the nature of the borrowing process was not evaluated favorably by study respondents. Too often, these findings show, consumer behavior professionals have not assessed the assumptions and relevance of social science constructs in terms of a market behavior context. In addition, the unidirectional nature of the borrowing process suggests that consumer behaviorists have not successfully integrated knowledge from allied fields in a manner that promotes contributions from consumer behavior to such other disciplines. However, respondents do believe that the study of consumer behavior has immeasurably strengthened the foundations of marketing.

Respondents also recognized that many consumer behavior research practices have been of questionable quality in the key areas of appropriateness and relevance of borrowed constructs, and in the improper use of consumer behavior research findings. Respondents also believe that many research findings are invalid because of reliance on incomplete theories.

Finally, in spite of the past problems experienced in consumer behavior research, ACR members generally expressed optimism regarding future prospects for the development of more complex models and of criteria to evaluate these models.

Two caveats must be noted when assessing the results of this study. First, the authors exercised subjective judgment in selecting the three major issues examined in this study. Other students of consumer behavior may desire to examine different topics in future research efforts. Second, while the sample utilized in this study encompassed a group of individuals with consumer behavior interests, similar investigations using different samples may produce other findings about the status of consumer behavior. However, this study provides an empirically based perspective of the current status of the field within the context of these limitations and serves as a benchmark study for future research.

#### FOOTNOTES

1. Gary T. Ford and Philip G. Kuehl are Assistant Professor and Associate Professor of Marketing, respectively, College of Business and Management, University of Maryland. Robert F. Dyer is Assistant Professor of Marketing, School of Government and Business Administration, The George Washington University.



## REFERENCES

- Engel, J. F., Kollat, D. T. and Blackwell, R. E. Consumer behavior. New York: Holt, Rinehart and Winston, Inc., 1973.
- Howard, J. A. New directions in buyer behavior research. Proceedings, Fall Conference, American Marketing Association, 1971, 375-80.
- Kollat, D. T., Blackwell, R. D, and Engel, J. F. The current status of consumer behavior research: Developments during the 1968-1972 period. Proceedings, Annual Conference, Association for Consumer Research, 1972.
- Sheth, J. W. The future of buyer behavior theory. Proceedings, Annual Conference, Association for Consumer Research, 1972, 562-75.
- Tucker, W. T. Consumer research: Status and prospects. Proceedings, Winter Conference, American Marketing Association, 1967, 267-69.
- Ward, S. and Robertson, T. S. Consumer behavior: Theoretical sources. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973.

## APPENDIX

## VERIFICATION OF INTRA-TABLE STATEMENT HOMOGENEITY

In order to increase our confidence that the statements included in each of the three tables in this paper were in fact investigating common dimensions concerning consumer behavior, we employed the following analysis procedure. The table of eighteen different statements from all three tables were grouped together and a principal components factor analysis (BMD 03M) was run. All factors with an eigenvalue of greater than 1.00 were then rotated according to the varimax rotation procedure and the resulting factor loadings were examined. We believed that if statements from the same table belonged together, they would exhibit high factor loadings on the same factors. Conversely, if statements from different tables loaded highly on the same factors we would be less confident concerning intra-table statement homogeneity.

Following this analysis procedure, seven factors accounting for 58 percent of the total variance were extracted and rotated. The results of this procedure are given in Table 4 which shows for each factor all statements with a loading of at least  $\pm .35$ . For all factors the statements showing the highest loadings are from the same table, therefore, lending support to our grouping of Likert statements within tables.

TABLE 4

## Statements With Highest Loadings on Each Factor\*

	Statement Number	Statement Number	Statement Number	Statement Number
Factor 1	3-5 (.808)	3-6 (.807)	3-7 (.702)	
Factor 2	2-3 (.750)	2-2 (.622)	2-1 (.602)	2-4 (.426)
Factor 3	1-3 (.781)	1-5 (.685)	1-4 (.365)	
Factor 4	2-6 (.699)	2-5 (.548)	2-4 (.377)	
Factor 5	3-2 (.744)	3-4 (.619)		
Factor 6	3-3 (.852)	3-1 (.420)		
Factor 7	1-1 (.751)	1-2 (.716)		

\*Factor loadings are shown in parentheses.



## OBSERVING STAGES IN CONSUMER DECISION PROCESSES

Johan Arndt<sup>1</sup>The Norwegian School of Economics  
and Business AdministrationEdgar Crane<sup>2</sup>

University of Houston

One-hundred husband-wife dyads were asked to determine among themselves how a hypothetical monetary gift was to be spent. The discussions were observed and tape-recorded by field workers. The messages exchanged were coded by the Bales Interaction Process Analysis system. The results showed some tendency for the interaction to change from a main emphasis on solving problems of orientation, to problems of evaluation, and subsequently to problems of control. Over the decision process, the relative importance of positive reactions increased.

Most current formulations of consumer decision making postulate that consumers go through a specified sequence of mental and behavioral steps in their decision processes. (Engel, Kollat, and Blackwell, 1973; Howard and Sheth, 1969; and Robertson, 1971.) A common feature of many of the formulations is that they build on hierarchy-of-effects models claiming that on their way from ignorance of a product to purchase, consumers move in one direction through a given sequence of cognitive, affective, and conative stages.

While such models at first sight seem to have an appealing intuitive validity, there is so far actually little direct empirical evidence that they adequately describe typical purchase decisions. Some of the studies purporting to support the hierarchy-of-effects model are interview studies in which the researcher seems to have induced respondents to accept the model (before being able to answer questions about communications behavior in the various stages), while other studies using less structured methods may have been subject to response bias as respondents may have attempted to rationalize their decision processes. Palda who critically reviewed the literature in this area, including survey studies as well as longitudinal approaches and field experiments, found little evidence that positive changes in either cognitive or affective states preceded purchase (Palda, 1966).

From a theoretical viewpoint, the hierarchy-of-effects model may be too narrow in scope, first by dealing with individual rather than family or other group purchase decisions and, second, by being stimulus oriented rather than oriented toward the consumer problem solving process.

The study to be reported here was undertaken to find whether small group theory and method may contribute to a better understanding of the anatomy of intrafamilial decision processes on spending matters.

In the behavioral sciences, a large body of theory and research findings on small group processes is available. A substantial part of this literature is concerned with the interaction in a group involved in developing group

solutions to various types of problems (Collins and Raven, 1969; and Kelley and Thibaut, 1969). A priori one might think that this literature should be particularly important for consumer behavior studies, particularly those focusing on group behavior such as is the case in intrafamilial decision making. However, so far small group theory seems to be neglected in the marketing literature. An exception is Davis and Silk (1971).

In the present study, both the theoretical framework and the method utilized have been developed by Bales.

### The Bales Approach

In a variety of research settings, Bales and his colleagues have found that the problem solving processes of small groups are more effective if some prescribed order were followed. More specifically, in so-called full-fledged problem solving, the Bales group has found that decision makers in their interaction move qualitatively from a relative emphasis upon problems of orientation ("what is it"), to problems of evaluation ("how do we feel about it") and subsequently to problems of control ("what shall we do about it"), and that the relative frequencies of both positive and negative reactions tend to increase toward the end of the process (Bales and Strodtbeck, 1951). It is assumed that this pattern is natural because earlier stages are "functionally prerequisite" to later ones.

Bales has developed a method for studying intragroup interaction referred to as Interaction Process Analysis. The heart of the method is that the interaction is observed, and that the individual interaction acts are recorded and classified by a 12-category system (Bales, 1950). A simplified statement of the system is shown in Table 1. In this scheme, categories 6 and 7 are viewed as dealing with problems of orientation; 5 and 8 deal with problems of evaluation; 4 and 9 with problems of control; 1, 2, and 3 with positive reactions; and finally 10, 11, and 12 with negative reactions.

An observational method such as the Bales approach does not rely on respondents' reconstructions of their decision processes or is subject to bias because the researcher superimposes a scheme of stages on the respondents. Instead, the method will detect the natural tendencies that exist in the data within the constraints given by the research setting and the classification system for the interaction acts.

### Method<sup>3</sup>

The field approach utilized is an extension of an approach earlier applied by Kenkel, who studied intrafamilial decision processes in the U.S. around 1960 (Kenkel, 1957; Kenkel, 1961a; and Kenkel, 1961b).

One-hundred husband-wife dyads selected by random sampling from public records in five electoral districts in Bergen, Norway, were asked to assume that they had received a monetary gift<sup>4</sup> and were then asked to determine between themselves just how this money should be spent, with the stipulation that the amount could not be saved. Each interaction process was observed by a field worker who made a tape recording of the discussion. The messages exchanged were coded by the Bales system. In addition, some background information about the respondents was obtained.

TABLE 1  
 Definition of Interaction Categories  
 and Distribution of Interaction Acts by Category

Category		Per Cent
Social-Emotional Area: Positive Reactions	1. <u>Shows solidarity</u> , raises other's status, gives help, reward	1.5
	2. <u>Shows tension release</u> , jokes, laughs, shows satisfaction	4.1
	3. <u>Agrees</u> , shows passive acceptance, understands, concurs, complies	9.7
Task Area: Attempted Answers	4. <u>Gives suggestion</u> , direction, implying autonomy for other	11.4
	5. <u>Gives opinion</u> , evaluation, analysis, expresses feeling, wish	18.1
	6. <u>Gives orientation</u> , information, repeats, clarifies, confirms	35.4
Task Area: Questions	7. <u>Asks for orientation</u> , information, repetition, confirmation	7.7
	8. <u>Asks for opinion</u> , evaluation, analysis, expression of feeling	4.9
	9. <u>Asks for suggestion</u> , direction, possible ways of action	2.1
Social-Emotional Area: Negative Reactions	10. <u>Disagrees</u> , shows passive rejection, formality, withholds help	3.3
	11. <u>Shows tension</u> , asks for help, withdraws out of field	1.1
	12. <u>Shows antagonism</u> , deflates other's status, defends or asserts self	.7
Total Base		100.0 (3,044)

Findings<sup>5</sup>

The average discussion lasted for only 3 minutes and 36 seconds, suggesting a decision process very compressed in time. A total of 3,044 interaction acts (including answers from the field workers to direct questions from husbands or wives) were recorded. The mean interaction rate of between 8 and 9 acts per minute found is in line with findings reported by Willett and Pennington who used the Bales scheme to analyze customer-salesman interactions in retail stores (Willett and Pennington, 1966, pp. 606-607), while smaller than the rate reported by Bales for problem solving groups - 15 to 20 acts per minute (Bales, 1958, p. 438). The relative frequencies shown in Table 1, with predominance of categories 5 and 6, are similar to patterns reported by Bales (1958, pp. 438-439) and Willett and Pennington (1966, pp. 607-608).

Let us now return to the question of stages in the decision processes as proposed by Bales.

To test the Bales hypothesis, first, the cycle of operations (or the number of interaction acts) in each discussion was divided into three thirds so as to produce a first, second, and third phase. This was in this case roughly equivalent to a division based on time. Second, the relative number of the five types of interaction categories was computed. The results which are presented in Table 2 show a significant departure from chance

TABLE 2  
Relative Frequency of Acts by Type and Phase

Type	Phase			Total
	First	Second	Third	
Orientation	48%	40%	42%	43%
Evaluation	21	25	23	23
Control	14	16	11	14
Positive Reactions	12	14	20	15
Negative Reactions	6	5	4	5
Total	101%	100%	100%	100%
Base	(1,016)	(1,014)	(1,014)	(3,044)
	$\chi^2 = 43.9$	$p < .001$	(8 d.f.)	

expectations. In other words, there was a relationship between phase and relative importance of interaction type. However, all relationships were fairly weak, and in several cases the relationships were in a direction other

than expected. Nevertheless, the main pattern in Table 2 seems to give at least some support to the Bales hypothesis. Hence, orientation acts culminated in the first phase and evaluation acts in the second, while the relative importance of positive reactions increased over the time cycle. Control acts culminated unexpectedly in the second and not in the third phase. It is interesting to note that toward the end of the cycle, negative acts tended to be displaced by positive ones. In this respect, our results parallel findings of Willett and Pennington for customer-salesman interactions resulting in a transaction (Willett and Pennington, 1966, pp. 614-616).

A possible explanation for the somewhat weak results may lie in the fact that the average dyad considered about 4 alternative items, of which about one-half was accepted. There may have been tendencies for all three phases to emerge for each alternative considered.

We then devised a second test relating to the time sequence of acts by type. In Table 3 where the results are shown, positive and negative reactions have been combined as no prior hypothesis existed as to which one was expected to precede the other.

TABLE 3  
Relative Frequencies of Acts by Type  
and by Type of Succeeding Act

Type of Succeeding Act	Type of Act				Total
	Orient-ation	Evalu-ation	Control	Positive or Negative Reactions	
Orientation	51%	35%	29%	39%	42%
Evaluation	21	26	16	25	22
Control	12	13	15	16	13
Positive or Negative Reactions	14	24	38	13	19
No Succeeding Act <sup>x)</sup>	3	2	1	6	3
Total	101%	100%	99%	99%	99%
Base	(1,314)	(700)	(411)	(619)	(3,044)

<sup>x)</sup> Not included in the Chi Square analysis

$$X^2 = 178.7 \quad p < .001 \quad (9 \text{ d.f.})$$



As suggested by the large Chi Square value, there was a relationship between act and type of preceding act. Orientation acts tended to be followed by other orientation acts, while evaluation acts were particularly likely to be followed by orientation acts or other evaluation acts. The likelihood that the succeeding act was concerned with control problems increased when the type of act changed from orientation, through evaluation, to control. A similar relationship was found for succeeding acts of the positive or negative reactions type. For instance, as many as 38 per cent of the acts following control acts were of this type. A possible explanation is that control acts impair the social-emotional relations in the group and disturb the equilibrium and hence create a need for social-emotional acts to reduce tension.

While the relationships in Table 3 were fairly weak as was the case in Table 2, the overall pattern seems to be to some extent consistent with the Bales hypothesis.

#### Concluding Comments

This exploratory study of husband-wife decision making has used a theoretical framework and a method developed by Bales. The results give some support to Bales' hypothesis stating that the process tends to move qualitatively from a relative emphasis on solving problems of orientation, to attempts to solve problems of evaluation, and subsequently to attempts to deal with problems of control. Toward the end of the decision process, positive reactions were found to increase in importance.

The relationships found, however, were not strong, and there were also several deviations from the theoretical expectations, suggesting some "noise" in the data. Nevertheless, the results seem interesting enough to warrant further work in this area in order to search for "natural" stages in inter-familial or other collective decision making.

#### FOOTNOTES

1. Professor of Business Administration, the Norwegian School of Economics and Business Administration.
2. Professor of Marketing, University of Houston, and Visiting Professor at the Norwegian School of Economics and Business Administration at the time when this study was undertaken.
3. Appreciation is extended to the following students who participated in planning and implementing the study: Hans A. Brandsnes, Bent Bugge, Arnt Buvik, Kjell E. Grude, Inge Rylandsholm, Åge H. Salin, Vemund P. Smørðal, and Agnar S. Tjeldnes.
4. Fifty couples "received" an amount of N.kr. 500 (about \$ 90), while the amount for the other 50 dyads was N.kr. 3,000 (about \$ 550). In the part of the study to be reported here, however, we will not be concerned with this experimental variation.

5. Other findings from this study are reported in Johan Arndt and Edgar Crane, Marital roles in intrafamilial decision making on spending matters, paper presented at the 1974 Marketing Educators' Conference of the American Marketing Association in Portland, Oregon.

## REFERENCES

- Bales, R.F. Interaction process analysis. Cambridge: Addison-Wesley Press, Inc., 1950.
- Bales, R.F. Task roles and social roles in problem-solving groups. In E.E. Maccoby, T.M. Newcomb, & E.L. Hartley (Eds.), Readings in social psychology. New York: Holt, Rinehart & Winston, Inc., 1958, 437-447.
- Bales, R.F. & Strodtbeck, F.L. Phases in group problem solving. Journal of Abnormal and Social Psychology, 1951, 46, 485-495.
- Collins, B.E. & Raven, B.H. Group structure: Attraction, coalitions, communication, and power. In G. Lindzey & E. Aronson (Eds.), The handbook of social psychology. (2nd ed.) Vol. IV. Reading: Addison-Wesley Publishing Company, Inc., 1969, 102-204.
- Davis, H.L. & Silk, A.J. Small group theory and marketing research. Working paper No. 548-71. Cambridge: Alfred P. Sloan School of Management, Massachusetts Institute of Technology, 1971.
- Engel, J.F., Kollat, D.T., & Blackwell, R.D. Consumer behavior. (2nd ed.) New York: Holt, Rinehart & Winston, Inc., 1973.
- Howard, J.A. & Sheth, J.N. The theory of buyer behavior. New York: John Wiley & Sons, Inc., 1969.
- Kelley, H.H. & Thibaut, J.W. Group problem solving. In G. Lindzey & E. Aronson (Eds.), The handbook of social psychology. (2nd ed.) Vol. IV. Reading: Addison-Wesley Publishing Company, Inc., 1969, 1-101.
- Kenkel, W.F. Family interaction in decision-making on spending. In N.N. Foote (Ed.), Household decision-making. Consumer behavior. Vol. IV. New York: New York University Press, 1961a, 140-164.
- Kenkel, W.F. Husband-wife interaction in decision making and decision choices. Journal of Social Psychology, 1961b, 54, 255-262.
- Kenkel, W.F. Influence differentiation in family decision making. Sociology and Social Research, 1957, 42, 18-25.
- Palda, K.S. The hypothesis of a hierarchy of effects: A partial evaluation. Journal of Marketing Research, 1966, 3, 13-24.
- Robertson, T.S. Innovative behavior and communication. New York: Holt, Rinehart & Winston, Inc., 1971.

Willett, R.P. & Pennington, A.L. Customer and salesman: The anatomy of choice and influence in a retail setting. In R.M. Haas (Ed.), Science, technology, and marketing. Proceedings of the Fall 1966 Conference of the American Marketing Association. Chicago: American Marketing Association, 1966, 598-616.

FAMILY ROLE STRUCTURE AND HOUSING DECISIONS<sup>1</sup>

Donald J. Hempel<sup>2</sup>  
University of Connecticut

This study explores the determinants and effects of family role structures in house buying decisions. It extends the earlier work of the author by examining the relative predictive value of selected family characteristics which other studies have identified as important determinants of role structure. It deals with the question of whether it makes any difference how roles are allocated within the family by analyzing the correlation of role structure with decision process variables. Comparisons are presented of both husband and wife role perceptions in two different cultural settings.

## Introduction

Family roles in buying decisions have received a great deal of attention from consumer researchers in the past year (e.g., Ferber and Lee, 1974; Davis and Rigaux, 1974; Hempel, 1974; Cunningham and Green, 1974). The empirical basis for developing models of family buying behavior is expanding rapidly, and there have been several recent efforts to formulate a theoretical framework for incorporating this information (Sheth, 1974; Granbois, 1972). Much of the evidence is available only in fragments however, and there is need for a more integrated coverage of family buying decisions as a behavioral system. More specifically, empirical research is needed concerning the relationship among three sets of variables: (1) the patterns of individual responsibility and influence within the decision-making unit--i.e., role structures; (2) predictors of the role structures which are likely to operate in specific buying situations--i.e., determinants; and (3) implications of different role allocations for the economic behavior of the household--i.e., effects.

Despite the extensive literature dealing with family decision making, there are significant limitations to this information base. As researchers in other disciplines have noted (e.g., Safilios-Rothschild, 1970), the available research findings concerning family role differentiation are too often based upon simple measures of general influence derived from the perceptions of wives. Many of the studies dealing with determinants of role structure were conducted more than a decade ago. Some recent studies have raised concerns about the reliability of this evidence for developing current theories of family buying decisions (e.g., Cunningham and Green, 1974; Ferber and Nicosia, 1972; and Davis, 1971). The growing emphasis upon practical implications of consumer research has also increased sensitivity to the fact that relatively little is known about the relationship of marital roles to other decision process variables (Sheldon, 1973).

## Relevant Research On Family Decision Making

Much of the empirical research concerning family buying decisions has been summarized in a series of review articles published within the last three years (Ferber, 1973; Granbois, 1972; and Sheth, 1971). It is sufficient to note here some of the more recent findings concerning husband-wife influence in family housing decisions. Home buying decisions receive special attention because they are considered to be a theoretical extremity of the "degree of deliberation" continuum on which a typology of buyer behavior might be constructed.

There is substantial evidence that the perceptions of both husband and wives should be considered in efforts to explain or predict family buying decisions (Hempel, 1974; Ferber, 1973; Davis, 1970). In a recent study of the asset accumulation behavior of young couples, Ferber and Nicosia (1972) stress the need to consider both the separate influences of each spouse and how they interact. Their findings concerning real estate investments (including home ownership) indicate that husband and wife influences may operate in opposite directions, perhaps reflecting intrafamily conflict on spending priorities. For example, attitudes toward the priority of savings were significantly related to real estate investment behavior in all but one instance, but the regression coefficients for husband and wife had opposite signs in every instance!

One of the more interesting developments in recent studies of family buying decisions is the growing conviction that husband-wife roles vary with the type of decision and with the product category (Ferber and Nicosia, 1972; Davis, 1970; Jaffee and Senft, 1966). The marital role allocations by decision area for 25 different product categories were examined by Davis and Rigaux (1974) in their study of Belgium couples. They found that housing decisions were highly syncretic across all three stages of problem recognition, search, and final decision. Contrary to the other products considered, housing decisions did not become less specialized between the initial and final stages. They reported little variability in role consensus among the three phases of the decision process for most products, but again housing was an exception. The evidence presented shows relatively high intrafamily role consensus for housing decisions in the problem recognition and final decision stages, while agreement at the search stage was substantially lower (81%, 82%, and 60% respectively). These findings support treatment of housing as a special type of purchase decision in the development of a typology for marital role differentiation. The position of housing decisions at the center of the "syncretic" classification in Davis and Rigaux's feasibility triangle suggests that it might be useful to consider this product class as particularly representative of the role category.

The implications of role structure for buying behavior were considered in a recent study by Ferber and Lee (1974). They found that families who assign a high priority to savings in the form of real estate are more likely to allocate decision making responsibility (i.e., the "family financial officer" role) to one member rather than both. Their findings indicate that the probability of buying a house during the early years of marriage was greater when the husband was the dominant decision maker. This tendency toward specialization in the financial management roles among young home buyers may encourage some compensatory allocation of other roles in the home buying decision process.

Most of the empirical studies concerning role allocations in the purchase of major durables indicate that joint decision making is the modal case. One very recent study suggests that shared decisions are becoming more common for these product categories, perhaps as a consequence of major social changes (Cunningham and Green, 1974). As the summary of findings presented in Table 1 evidence, both of these generalizations appear to be appropriate for housing decisions. The apparent trend toward the merging of husband and wife roles may be affected by respondent efforts to present family role distributions which appear equitable in an environment of growing social concern for equality. Some evidence of a compensating pattern of role allocations was presented by Hempel (1974). Efforts to characterize a complicated decision process in a general statement of husband-wife influence may represent attitudes about how roles should be allocated as much as how the roles were actually distributed. To the extent that this response behavior is associated with a convergence of future purchase behavior toward the attitude, the net effect will be to improve the predictive value of the data. The measures of role structure therefore may have desirable reliability qualities even though they might be lacking validity.

Despite some possible confounding of attitude and behavior in the measures of role structure, they can be considered potentially useful bases for market segmentation. Evaluation of this potential requires more information about the linkage between role structure and more readily observable characteristics of the household. It also requires information about the implications of role structure for other aspects of consumer decision processes, such as search, purchase behavior, post-purchase satisfaction. The results presented here are a preliminary attempt to address these needs.

#### 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 considerably by nonreturns and the elimination of all observations with missing information. Details concerning the research design and questionnaires have been published elsewhere (Hempel, 1970).

#### Determinants of Family Role Structure

The bulk of research on the determinants of family role allocations in purchase decisions has focused upon the socioeconomic characteristics of the household. There are several general hypotheses which have been advanced to explain the extent of joint decision-making in buying decisions: (1) the greater the resources, contribution, and status of the individual, relative to his or her spouse, the greater the influence--e.g., education, occupational prestige, income contribution, and employment status; (2) more extensive experience as a decision-making unit is associated with a reduction in

TABLE 1

## Major Studies of Family Role Structure in Housing Decisions

	<u>Percent Distribution of Role Pattern</u>				
	Husband	Joint	Wife	N	Respondent(s)
1. Detroit Area, 1955 (Sharp and Mott, 1956)	18%	58%	24%	727	wives only
2. Detroit Area, 1955 (Blood and Wolfe, 1960)	18	58	24	731	wives only
3. Hartford Area, 1968 (Hempel, 1972)	20	67	13	152	husbands and wives separately
4. Vancouver B.C., 1969 (Kelly and Egan, 1969)	N.A.	"nearly two-thirds"	N.A.	64	husbands and wives separately
5. Northwest England, 1971 (Hempel, 1972)	17	74	9	196	husbands and wives separately
6. Louvain Beligum, 1971 (Davis and Rigaux, 1974)	N.A.	69	N.A.	73	husbands and wives separately
7. Houston Area, 1973 (Cunningham and Green, 1974)	12	77	11	248	wives only

Note: The following questions were used to develop the data base for each of the studies summarized above:

1. Who decides what house or apartment to take?
2. Who usually makes the final decision about what house or apartment to take.
3. Average of husband and wife responses concerning relative responsibility and influence for four decisions--neighborhood, style of house, when to purchase, and acceptable price.
4. "Decision to begin searching for a house".
5. Same as 3
6. Average of husband and wife responses concerning relative influence across three stages of decision process--problem recognition, search, and final decision.
7. Who decides the location, price, and size of the house or apartment to take?

joint-decisions--e.g., age, years married, later stages of family life cycle, and income; and (3) greater "connectedness" of the family's social network and social distance from the middle class are inversely related to the degree of joint-decision making. Discussions of the empirical studies which serve as a basis for these generalizations can be found in several recent publications (Ferber, 1973; Engel, Kollat, and Blackwell, 1973; and Granbois, 1972.

Table 2 shows the correlation between four measures of marital roles in housing decisions and several sets of determinants. The first two role measures are dichotomized variables which represent the classification scheme proposed by Herbst (1954) applied to the housing decision patterns described by Hempel (1974). The third role measure represents the extent of joint decision making as the total number of shared decisions reported by either spouse. The overall measure of the husband-wife influence perceived by each spouse was constructed by aggregating the decision scores (1=wife, 2=joint, 3=husband) over five decision areas (neighborhood, style, price, timing of purchase, and mortgage source). More elegant statistical results could be developed by calculating biserial correlations for the first two role structure measures, but the product moment correlations presented here provide a consistent and useful first approximation for evaluating the relative importance of different determinants.

If predictive value is used as the criterion for judging importance, education and occupational status appear to be the most significant determinants of role structure in house buying decisions. The best single predictor across role measures and markets was the ratio of husband/wife occupational status. This indicator of relative resources tended to be a better predictor of the role perceptions reported by husbands than those of wives. The husband's education also was closely related to the role perceptions of both spouses in Connecticut, but the educational variable was considerably less significant in England. It is noteworthy that the wife's occupational status appears to have much more influence upon the husband's perceptions of role structure than her own. The measures of the wife's employment status, particularly the number of hours worked, were also correlated with the husband's perceptions, but these variables were not significant determinants for the wives.

The direction of relationship was generally consistent with that found in previous studies. Age, years married, occupational status of the wife, and her employment status were all inversely related to husband dominance of the buying decisions. Education, the husband's occupational status, and family income were positively correlated with the measures of husband dominance. The curvilinear hypothesis that joint decision making is less common in the lower and upper social classes was not confirmed for the lower class. This apparent discrepancy can be reconciled with the recognition that few members of the lower class are likely to appear in a sample of home buyers. The "lower" educational and occupational groups in this study can more appropriately be regarded as members of the lower-middle social classes.

In general, the husband dominant patterns can be predicted more readily from the determinants considered here than either the syncretic classification or the total number of joint decisions. A multiple regression analysis of the different role structure measures indicates that five or six of the determinants can "explain" about 25% of the variance in either dominance measure, whereas they will account for less than 20% of variance in the other two measures. In either case, there is need for consideration of other determinants of role structure in addition to those presented here.



TABLE 2

Possible Determinants of Family Role Structure in House Buying Decisions  
Among Connecticut (N=137) and English (N=106) Households

Determinant	Correlation Coefficients for Measures of Husband/Wife Role Perceptions							
	Role Structure Classification				Total Number of		Overall Dominance Index	
	Husband Dominant		Syncratic		Joint Decisions (H+W)		(Sum 5 Decision Scores)	
	HTC	NWE	HTC	NWE	HTC	NWE	HTC	NWE
<u>Husband-Wife Demographics</u>								
Husband's age (years)	- 8/1	-12/16	1/5	- 2/2	- 1	- 3	-13/6	-21/14
Wife's age (years)	-11/6	-10/13	3/3	- 4/3	1	- 6	-17/11	-18/14
Years Married	-18/15	- 4/9	7/5	- 8/7	7	-11	-26/19	-15/6
Husband's education (yrs)	34/26	15/12	-25/14	- 4/8	-21	- 4	22/31	15/17
Wife's education (yrs)	19/1	6/10	- 9/0	10/-6	- 8	14	11/8	15/9
H's occup. status (scale 3-12)	27/22	17/21	-14/0	- 8/2	-10	-11	16/25	10/31
W's occup. status (scale 3-12)	-16/6	-10/8	21/0	20/19	6	8	-28/8	- 9/6
<u>Household Characteristics</u>								
FLC1-young, no dep. (0,1)	0/2	5/10	3/-9	5/2	6	- 8	1/3	14/18
FLC2-young, dep. (0,1)	25/13	2/10	-13/11	2/-3	- 3	12	10/-1	6/-7
SCL1-High educ. & occup. (0,1)	26/20	16/19	-10/9	- 5/16	-11	-18	13/20	12/24
SCL2-low educ. & occup. (0,1)	-32/18	-16/16	16/1	3/5	12	8	-19/28	- 9/30
No. persons in household	6/3	11/1	- 5/+9	- 0/7	0	11	6/5	- 3/+1
Annual family income (\$ or £)	11/17	5/5	- 5/24	2/0	-21	5	5/23	14/18
<u>Relative Resources</u>								
Age (H's age + W's age)	9/14	- 2/11	-12/7	2/3	- 9	5	9/10	- 2/1
Education (H's + W's)	16/27	6/6	-16/16	-11/5	-16	-21	12/24	- 1/+12
Occup. Status (H's + W's)	35/18	21/24	-26/1	-20/12	-14	-14	35/22	13/32
<u>Wife's Employment Status</u>								
Hours empl. outside home	-21/3	-17/2	29/1	19/8	18	11	-29/6	- 6/+6
Years empl. by firm	-23/14	-10/2	18/2	9/3	16	10	-27/17	-12/6
W's contrib. to income (%)	-12/8	-13/2	15/-3	11/3	14	10	-24/11	- 3/+3

Note: The figures in each cell represent the Pearsonian bivariate product moment correlation between the measure of role allocations for husband/wife and the variable listed as a determinant. These coefficients are shown without decimals to simplify the tabulation. The sign of the coefficients for both husband and wife are the same except in the few instances where the opposite is shown for the wife below the slash line.

The selection of appropriate role structure measurements is complicated by the finding that the predictive value of the determinants appears to vary by culture and sex of the respondent. In Connecticut, the family characteristics generally predict the husband's perceptions of role structure much better. The relative importance of determinants varies less by sex in England, but the role structure measures also are less predictable. Perhaps this pattern reflects a cultural setting in which the marital role perceptions and expectations are more independent of the socioeconomic characteristics which often are used in our culture to characterize the family as a decision making unit.

### Decision Process Implications of Role Structure

Does the pattern of role allocations within family buying decisions influence other aspects of buyer behavior? This question has not been investigated in much detail, but one recent study by Ferber and Lee (1974) suggests that such effects may be appreciable. The linkages of role structure to other behavioral variables should be a basic concern of both the pragmatist involved in strategy formulation and the theoretician oriented toward model building.

Table 3 presents the correlation of the role structures perceived by husbands and wives with measures of search behavior, expenditures and post-purchase satisfaction. These results indicate that the main effects of role allocations in housing decisions may show up in the interrelationships with expenditures for other goods. The other decision process variables were related to role structure in some cases, but the correlations were much less significant.

The amount which the family spent for complementary goods (furniture, appliances, decorating, etc.) during the two month period centered around the occupancy of their house was related to all of the role structure measures for English households. In Connecticut, this expenditure behavior emerges only in the wife dominant and syncratic households. A prevalence of joint decision making was associated with lower expenditures for complementary goods in both countries, as was the wife dominant case in England. Husband dominance and autonomic role allocations were directly related to the level of expenditures among English families. This pattern suggests that the wife's participation in house buying decisions may generate some tradeoffs in the buying decisions for other products. More role specialization in housing decisions may occur because the family is concurrently engaged in buying decisions for other major products.

The time required for the decision process (i.e., the decision span) was positively correlated with the role perceptions of both spouses in the husband dominant English households, but only with the wives' perceptions in similar Connecticut households. This pattern is repeated for the syncratic households, but the direction of relationship is negative. These results seem to suggest that when the wife is jointly involved in the house buying decisions, less time is required for the decision process. The impact of marital roles upon the extent of search (i.e., the number of houses inspected) appears to be negligible in all but one case--wife dominant English households were likely to examine more housing alternatives.

Satisfaction with the house purchased was not consistently related to family role structure. There is some evidence that family satisfaction is

TABLE 3

Correlation of Role Structure Measures With Selected Decision Process Variables

Role Structure	Decision Process Variables for Connecticut (N=151) and English (N=178) Households											
	Decision Span (mos)		No. of Houses Inspected		Price Paid Per Room		Expenditures for Complementary Goods		Index of Family Satisfaction			
	HTC	NWE	HTC	NWE	HTC	NWE	HTC	NWE	HTC	NWE	HTC	NWE
Husband Dominant												
H's perceptions	-.02	.15	.10	-.06	.13	.08	.07	.31	-.14	.01		
W's perceptions	.13	.13	.01	-.09	.06	.21	-.01	.35	.09	-.03		
Wife Dominant												
H's perceptions	-.06	.03	-.07	.16	-.04	-.10	.17	-.15	.06	-.14		
W's perceptions	-.05	.06	-.04	.21	-.04	-.17	.17	-.14	.03	-.15		
Syncratic												
H's perceptions	.04	-.17	-.03	-.07	-.07	-.10	-.16	-.27	.01	.03		
W's perceptions	-.15	-.15	-.02	-.04	-.04	-.13	-.20	-.21	.04	.11		
Autonomic												
H's perceptions	-.01	-.16	-.02	.04	-.04	.13	.06	.36	.11	.08		
W's perceptions	.07	.02	.06	-.03	-.03	.13	-.05	.19	-.23	.02		
Dominance Index												
H's perceptions	-.01	.10	.12	.08	.11	.11	.05	.21	-.08	-.04		
W's perceptions	.03	.08	.02	.06	.02	.28	-.05	.29	-.10	-.02		

Note: The figures on each cell represent the Pearsonian product moment correlation coefficients for the bivariate association between the role structure measures and decision process variables.

greater if neither spouse is perceived as dominant. The perception of role specialization (i.e., autonomic structure) among Connecticut wives was associated with lower levels of postpurchase satisfaction.

### Conclusions

The results of this study indicate that role structure in family house buying decisions is affected by household characteristics which are often used as bases for market segmentation. Education, occupational prestige, and the wife's employment status appear to be the most useful predictors. There was considerable variation in the relative importance of different determinants across sex, culture, and role structure measure.

The family characteristics and other influences upon role allocations considered here are an insufficient set of role structure determinants. Other variables such as attitude, experience, and decision making constraints (e.g., time pressure) should be incorporated in future studies to provide a more complete basis for models of family buying decisions. Measures of the attitude structure and personal involvement of each spouse, and their interplay with product experience variables appear to be particularly attractive areas for further research on housing decisions.

The evidence concerning role structure effects was rather disappointing. Though empirical research concerning these relationships is very scanty, the available conceptual models suggest a stronger association with search behavior and post-purchase satisfaction than the data indicate. These results suggest that role structure has relatively little direct influence upon other types of behavior within the decision process. The findings about the interrelationships between roles in housing decisions and expenditures for related goods are more promising, and they will be examined further for insights into the interdependency among buying decision processes.

### FOOTNOTES

1. This research was partially supported by funds from the University of Connecticut Research Foundation and the Marketing Department at the University of Lancaster, England. This assistance is gratefully acknowledged.
2. Donald J. Hempel is Professor of Marketing and Acting Head of the Marketing Department at the University of Connecticut.

### REFERENCES

- Blood, R.O. & Wolfe, D.M. Husbands and wives: The dynamics of married living. Chicago: The Free Press, 1960.
- Cunningham, I. C. M. & Green, R. T. Purchasing roles in the U.S. family, 1955 and 1973. Journal of Marketing, 1974, 38(4), 61-64.
- Davis, H. L. Measurement of husband-wife influence in consumer purchase decisions. Journal of Marketing Research, 1971, 8, 305-312.

- Davis, H. L. Dimensions of marital roles in consumer decision making. Journal of Marketing Research, 1970, 7, 168-77.
- Davis, H. L. & Rigaux, B. P. Perception of marital roles in decision processes. Journal of Consumer Research, 1974, 1(1), 51-62.
- Engel, J. F., Kollat, D. T., & Blackwell, R. O. Consumer behavior. (2nd Ed.) New York: Holt, Rinehart, 1973.
- Ferber, R. Family decision making and economic behavior. In E. B. Sheldon (Ed.), Family economic behavior: Problems and prospects. Philadelphia: J. B. Lippincott Co., 1973.
- Ferber, R. & Nicosia, F. Newly married couples and their asset accumulation decisions. In B. Strumpel, J. Morgan & E. Zahn (Eds.), Human behavior in economic affairs. San Francisco: Jossey-Bass, 1972.
- Ferber, R. & Lee, L. C. Husband-wife influence in family purchasing behavior. Journal of Consumer Research, 1974, 1(1), 43-50.
- Granbois, D. H. Decision processes for major durable goods. In G. Fisk (Ed.), New essays in marketing theory. Boston: Allyn and Bacon, 1972.
- Hempel, D. J. A comparative study of the home buying process in two Connecticut housing markets. Storrs, Connecticut: Center for Real Estate and Urban Economic Studies, The University of Connecticut, 1970.
- Hempel, D. J. Family buying decisions: A cross cultural perspective. Journal of Marketing Research, 1974, 11, 295-302.
- Jaffee, L. J. & Senft, H. The roles of husbands and wives in purchasing decisions. In L. Adler & I. Crespi (Eds.), Attitude research at sea. Chicago: American Marketing Association, 1966.
- Sheldon, E. G. (Ed.), Family economic behavior: Problems and prospects. Philadelphia: J. G. Lippincott Co., 1973.
- Sheth, J. N. A theory of family buying decisions. In P. Pellemans (Ed.), Insights in consumer and market behavior. Belgium: Namus University, 1971.
- Sheth, J. N. Models of buyer behavior: Conceptual quantitative, and empirical. New York: Harper & Row, 1974.

EFFECTS OF PRIOR DECISION-MAKING,  
 DEMOGRAPHICS, AND PSYCHOGRAPHICS ON  
 MARITAL ROLES FOR PURCHASING DURABLES<sup>1</sup>

Arch G. Woodside<sup>2</sup>  
 University of South Carolina

Can marital roles be accurately predicted for who actually makes the purchases of automobiles and rugs/carpets from information on marital roles in prior stages in the decision-making process and from demographic and psychographic data? What variables are most associated with marital roles of actual product purchasing? Results of this study of 200 husbands and 200 wives suggest that patterns for developing family types in consumer decision-making do exist.

Ferber (1973) has made a strong case for the greater need for product purchase models compared with brand purchase models:

. . . marketing researchers have been placing too much emphasis on using concepts from attitude theory to explain brand purchases and too little to explain product purchases. After all, the key problem is to anticipate sales of the product, which for many products may vary substantially from one year to another, whereas for most products brand shares vary little from one year or quarter to another.

This greater need for theoretical and empirical analysis of product decision-making by consumers may become apparent with the arrival of the 1974-1975 recession. Mid-October new car sales in 1974 were the lowest since 1964 and nearly 30% below 1973 sales (UPI, 1974), and, the 1974 slackening demand for other consumer goods suggests the basic question for consumer research is how consumers enter, pass through, and actually make product purchase decisions.

Fortunately, the study of family buying decisions has been concentrated on product purchase behavior (Hempel, 1974; Davis, 1970; Ferber and Lee, 1974; Davis and Rigaux, 1974; Sheth, 1974) which has produced substantial insights into the relative roles of husbands and wives in buying particular products and the effects of life cycle, social class, employment status of wife, social networks, and prior decision-making on these buying roles (Sheth, 1974).

The Engel, Kollat, and Blackwell (1973, p. 58) model of consumer behavior is the most related theoretical statement to extant research into family buying decisions and product purchase behavior. Both the Engel, Kollat, and Blackwell model and family buying research concentrate on the stages in cognitive decision processes of product purchases: problem-recognition, search for information, evaluation, store choice and shopping decisions, purchasing, and post-purchase processes. Other models and research of consumer behavior have concentrated on the relations between perceptions, attitude, intentions and purchases of brands (Howard and Sheth, 1969; Hansen, 1973).

The focus of the present study is the effects of prior decision-making, demographics, and psychographics on husband and wife roles for purchasing

durables. Can marital roles be accurately predicted for who actually makes the purchases of automobiles and rugs/carpets from information on marital roles in prior stages in the decision process and from demographic and psychographic data? What variables are most associated with marital roles of actual product purchasing? What are the differences in these salient variables for husband, wife, and syncretic purchases?

A basis for developing family types based upon the effects of prior decision-making and demographic and psychographic information would exist given substantially accurate predictions of marital roles in product purchasing. Davis (1970) has found that little basis for developing family types exists from patterns of husband-wife influence across several decisions when the decisions are not considered in independent and dependent relationships, e.g., where to buy not made dependent on when to buy. However, the relative influence of husbands and wives in making actual purchases may indeed be dependent upon prior decision-making (and demographic and psychographic variables) which would be suggested by the Engel, Kollat, and Blackwell (1973) model.

#### Method

The data were obtained from a cross-sectional survey of 200 families from three housing subdivisions in the Columbia, South Carolina, metropolitan area in 1972. Families were selected from a random sample of street tracts of single family dwellings with two families surveyed on each street. The social classes (defined by occupation and income ratings) were distinctly different for the families from the three subdivisions: upper-lower, lower-middle, and upper-middle social divisions.

Husbands and wives were interviewed separately in the 200 families. Interviewers were male and female graduate students. One interviewer went into each house separately. Families selected for the study were contacted prior to the interview by letter and telephone.

The survey instrument consisted of two parts. Husbands and wives answered a series of questions on the relative influence of each spouse for eight products: automobiles, lawnmowers, automatic washing machines, beer, rugs/carpets, cheese, television sets, and gardening supplies. Secondly, the husbands and wives completed a demographic and psychographic instrument which included questions on family life cycle, occupation of family head, wife employment, education, income, number of visits by friends in neighborhood, popularity, conservatism, club activities, advertising attitudes, opinion leadership, and other psychographic questions. The psychographic questions were selected from 300 factor analyzed questions developed by researchers at Purdue University (Tigert, 1969). Typically, four questions were included from each factor within the survey. Responses were true and false for each item and the factor was scored from 1 to 4.

The responses for each spouse on relative influence were scored by assigning a value of 1 when the husband's role was perceived as dominant, 2 for joint decisions, and 3 when the wife's role was dominant.

The decision-making questions for the eight products included questions on who brought up the idea of purchasing the product, discussion of the purchase with friends, neighbors, relatives, obtaining information from mass media, obtaining information from stores (dealers), style or type decisions, who visited stores or dealer showrooms, who selected the specific retail outlet to

purchase, who made the actual purchase, and who experienced dissatisfaction, if anyone.

The assumptions and limitations of the direct questioning approach of family roles in decision-making have been discussed elsewhere (Kenkel, 1961; Davis and Rigaux, 1974). In particular, the question of ability to recall with accuracy how the influence was distributed in some past decision-making session should be empirically studied. However, the direct questioning approach does permit a richness in data analysis since a large number of decision stages and products can be included in the same study and the validity of the measurement of perceived influence of husbands and wives separately has been supported (Davis, 1971). Engel, Kollat, and Blackwell (1973, p. 203) offer the following conclusion on measuring family role structure:

At the present time there does not appear to be any research design that can overcome the problems involved in identifying roles. The best interim approach seems to be to use direct questions about specific decisions and activities at each stage in the decision-making process for the marketer's product or service.

The following analyses of patterns of relative influence of husbands and wives are for two durable goods purchases, an automobile and rugs/carpets. Both of these products involve substantial financial outlays, extended periods of ownership, social importance, use by several family members, and most likely different patterns of husband and wife influence. Prior research (Cunningham and Green, 1974; Davis, 1970) has found husbands to be more often dominant than wives for some automobile decisions, e.g., what make of automobile to buy. Wives have been found (Jaffee Associates, 1965) to be more often dominant than husbands for some carpets/rugs decisions, e.g., who first brought up the idea and who actually purchased the product.

### Findings

Substantial variability between decisions and between products in husband-wife roles can be found in the distribution of husbands' and wives' responses as shown in Table 1. Both husbands and wives responses show high percentages of husband-dominant families for bringing-up the idea to purchase an automobile (70.6% of both husband and wife responses) and in actually purchasing the automobile (67% and 66.5% of the respective husband and wife responses). Some of the other decisions and purchase activities show decreases in husband-dominance and increases in syncretic behavior for automobile purchases, e.g., 66.7% of the husbands and 67.6% of the wives reported that visiting dealer showrooms to be a joint activity.

More husbands and wives reported that the wife had greater dominance in rugs/carpets decisions compared with automobile decisions. The percentage of wife-dominant families decreases from 60.1 for bringing-up the idea to purchase to 26.8% for actually purchasing among husband responses for rugs/carpets. Wife responses are similar in percent compared with husband responses. Higher percentages of husband-dominance is shown in Table 1 for making the actual purchase than for any other decision or activity, over 20% of both husbands and wives reported that the husband alone made the actual purchase for rugs/carpets.

High levels of agreement in relative influence among husbands and wives are shown in Table 1 (for aggregated data). The rank orders of percentages



TABLE 1  
Perceived Marital Roles of Husbands and Wives in Purchasing  
Automobiles and Rugs/Carpets in Percent (n = 200)

	Automobiles						Rugs/Carpets					
	Husband responses			Wife responses			Husband responses			Wife responses		
	H <sup>a</sup>	J <sup>b</sup>	W <sup>c</sup>	H	J	W	H	J	W	H	J	W
Brought up idea	70.6	18.0	11.0	70.6	20.6	8.8	9.1	30.8	60.1	10.2	33.0	56.9
Style	54.8	40.6	4.6	56.7	37.1	6.2	5.1	39.9	55.1	7.1	41.1	51.8
Size	58.9	36.5	4.6	56.9	38.5	4.6	10.6	46.0	43.4	13.8	45.4	40.8
Brand	59.4	38.1	2.5	60.4	35.9	3.6	9.6	45.7	44.7	14.8	44.9	40.3
How much to spend	51.8	45.2	3.0	55.0	41.3	3.7	19.2	55.1	25.8	20.7	53.0	26.3
Consulted friends	57.7	39.4	2.8	45.5	46.8	7.8	14.5	47.4	38.2	11.8	44.7	43.4
Consulted mass media	47.8	51.1	1.1	44.9	52.2	2.9	14.8	52.3	33.0	6.2	50.6	43.2
Consulted stores	59.9	39.5	0.7	56.3	43.0	0.7	13.6	55.3	31.1	16.7	51.5	31.8
Visited stores	33.3	66.7	0.0	31.9	67.6	0.5	5.2	67.0	27.8	6.7	62.6	30.8
Specific outlet	65.5	33.0	1.5	64.0	33.9	2.1	16.3	56.6	27.0	15.2	54.8	29.9
Actual purchase	67.0	31.4	1.5	66.5	31.4	2.1	21.6	51.5	26.8	24.6	47.2	28.2

<sup>a</sup>H = Husband    <sup>b</sup>J = Joint    <sup>c</sup>W = Wife

of families with husband-dominant or wife-dominant are nearly identical across the 11 decisions for automobiles and rugs/carpets. This finding has been previously observed for automobiles and furniture (Davis, 1970), and automatic washing machines and television sets (Woodside, 1972). Therefore, similar conclusions on relative influence would be likely reached by analyzing either aggregated husband or wife responses; at the same time, some substantial differences within families may exist for husband and wife responses.

Patterns of relative husband-wife influence across decisions and demographic-psychographic information for predicting who actually purchased the two products were analyzed through stepwise multiple discriminant analysis (MDA). Husbands and wives were separately grouped by husband, joint, or wife-dominance for actually purchasing automobiles and rugs/carpets. The independent variables included 10 prior decision-making variables as shown in Table 1 and 26 demographic-psychographic variables.

The average responses and standard deviations of four prior decision-making variables and one psychographic variable included in the first five variables to enter the multiple discriminant functions for automobiles among the husband responses are shown in Table 2. A total of 82.7% of the 191 husbands were correctly classified into husband or joint decision-making in actually purchasing automobiles. A wife-dominant group was not included in the analysis since so few wives dominated the actual purchase decision. Results are shown only for 5 variables in Table 2 since additional variables did not substantially increase the level of correctly classifying respondents.

The standardized betas shown in Table 2 indicate the most salient variables in discriminating relative husband-wife influence. The likelihood of jointly making the actual purchase of an automobile increases if friends (and neighbors or relatives) are jointly consulted prior to purchase. Jointly visiting showrooms and choosing a specific dealer increases the likelihood of jointly making the actual decision; the same finding holds for jointly deciding how much to spend for the automobile. Classification as husband-dominant in actually purchasing an automobile is more likely, the more favorable the attitude toward private brands.

Decisions on bringing-up the idea to purchase, style, size, brand, and consulting mass media were substantially less important than the variables discussed in correctly classifying husbands by relative influence in actually purchasing automobiles.

The analysis of the wife responses for automobiles is shown in Table 3. A total of 82.9% of the wives were correctly classified into husband-dominant or joint decision-making in making the actual purchase decision by 5 variables in the multiple discriminate functions.

High income level increased the likelihood of classification into husband-dominance in actually purchasing automobiles while high conservatism increased the likelihood of joint decision-making for the wife responses. Relative influences of husbands and wives in deciding how much to spend, visiting showrooms, and choosing a particular dealer were significantly related to the relative influence of husbands and wives in actually purchasing automobiles, as analyzed from the wife responses.

The percentages of correctly classified respondents for automobiles were substantially higher than would be expected by either the proportional chance criterion (57%) or the maximum chance criterion (68%), and these percent

TABLE 2

Means ( $\bar{X}$ ), Standard Deviations (s), Standardized Betas,  
and Confusion Matrix for Husband Responses  
for Automobiles and Five Variables

Variable	Relative influence			
	Husband (n = 130)		Joint (n = 61)	
	$\bar{X}$	s	$\bar{X}$	s
Pro private label brands	1.98	1.11	1.72	1.11
How much to spend	1.34	.51	1.84	.49
Consulted friends	1.09	.29	1.33	.54
Consulted stores	1.17	.38	1.56	.50
Specific outlet	1.15	.38	1.74	.44

Variable	Standardized betas	
	Husband	Joint
Pro private label brand	1.38	1.05
How much to spend	1.62	2.10
Consulted friends	1.46	3.26
Consulted stores	1.77	3.01
Specific outlet	1.54	2.99

Confusion matrix in percent			
Actual	Predicted		Total
	Husband	Joint	
Husband	84.6	15.4	130
Joint	21.3	78.7	61

TABLE 3

Means ( $\bar{X}$ ), Standard Deviations (s), Standardized Betas,  
and Confusion Matrix for Wife Responses  
for Automobiles and Five Variables

Variable	Relative influence			
	Husband (n = 127)		Joint (n = 66)	
	$\bar{X}$	s	$\bar{X}$	s
Income	4.46	1.13	4.10	.86
Conservatism	4.42	.87	3.77	1.37
How much to spend	1.26	.49	1.85	.40
Visited stores	1.56	.50	1.95	.22
Specific outlet	1.18	.39	1.72	.52

Variable	Standardized betas	
	Husband	Joint
Income	4.23	2.77
Conservatism	3.33	4.20
How much to spend	.92	1.62
Visited stores	2.35	1.30
Specific outlet	1.22	2.55

Confusion matrix in percent			
Actual	Predicted		Total
	Husband	Joint	
Husband	84.2	15.8	127
Joint	20.0	80.0	60

accuracies may be assumed to be impressive (Morrison, 1969). However, the results are somewhat biased upward since all the data were used to estimate the discriminate coefficients (Frank, Massy, and Morrison, 1965).

Data in Tables 4 and 5 are the means, standard deviations, standardized betas, and confusion matrixes for husband and wife responses in classifying families by relative influence in actually purchasing rugs/carpets. The relative influences on the size of the rug/carpet, visiting stores, and choosing the specific outlet were important in classifying families for both husband and wife responses. The greater the wife involvement in these decisions except for determining sizes, the greater the likelihood of wife-dominance in the actual purchase decision. Neighborhood popularity and consulting friends were significant in classifying relative influence in actually purchasing rugs/carpets for the husband responses, while the employment status of the wife and the relative influence on how much to spend were important in classifying relative influence in actually purchasing rugs/carpets for the wife responses.

Wife employment was entered into the functions as a dummy variable: 1 for wife as a full-time housewife and 2 for wife employed outside the home. Results in Table 5 would suggest that wife employment outside the home increases the likelihood that the wife would be involved in making the actual purchase of rugs/carpets.

The percentages of correct classifications for relative influence in actually purchasing rugs/carpets are substantial for joint and wife-dominance and less impressive for husband-dominance as shown in Tables 4 and 5.

#### Discussion

The findings support the hypothesis that some basis of developing family types does exist through the analysis of relative influence in prior decision-making and demographic-psychographic data. The predominance of relative influence in prior decision-making in predicting husband and wife roles in actual purchases of automobiles and rugs/carpets supports the cognitive-process model of purchase behavior by Engel, Kollat, and Blackwell (1973).

Demographic and psychographic variables do appear to be associated in a more limited role with actual purchase behavior compared with prior decision-making variables. Demographic and psychographic data may be more likely associated with the relative influences of husbands and wives in initiating the idea of purchasing rather than the purchase itself. This hypothesis should be explored in future research.

#### FOOTNOTES

1. Data collection assistance of Patty Richardson and Timothy Bice is gratefully acknowledged.
2. Associate Professor of Business Administration and Program Director of Marketing.

TABLE 4

Means ( $\bar{X}$ ), Standard Deviations (s), Standardized Betas,  
and Confusion Matrix for Husband Responses for  
Rugs/Carpets and Five Variables

Variable	Relative influence					
	Husband (n = 42)		Joint (n = 100)		Wife (n = 52)	
	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
Neighborhood popularity	2.81	.83	2.73	.69	2.88	.65
Size of carpet	2.12	.71	2.15	.61	2.83	.43
Consulted friends	1.90	.48	2.13	.39	2.17	.46
Visited stores	1.86	.47	2.06	.31	2.83	.38
Specific outlet	1.50	.55	1.99	.41	2.83	.43

Variable	Standardized betas		
	Husband	Joint	Wife
Neighborhood popularity	6.29	4.95	5.75
Size of carpet	3.64	2.69	2.43
Consulted friends	2.43	2.21	2.03
Visited stores	6.04	4.34	7.74
Specific outlet	3.34	3.48	5.28

Confusion matrix in percent				
Actual	Predicted			Total
	Husband	Joint	Wife	
Husband	64.3	30.9	4.8	42
Joint	9.0	87.0	4.0	100
Wife	1.9	7.7	90.4	52

TABLE 5

Means ( $\bar{X}$ ), Standard Deviations (s), Standardized Betas,  
and Confusion Matrix for Wife Responses for  
Rugs/Carpets and Five Variables

Variable	Relative influence					
	Husband (n = 48)		Joint (n = 92)		Wife (n = 55)	
	$\bar{X}$	s	X	s	$\bar{X}$	s
Wife employed	1.35	.48	1.43	.50	1.62	.49
Size of carpet	2.02	.73	2.08	.62	2.80	.45
How much to spend	1.58	.65	1.97	.50	2.62	.59
Visited stores	1.88	.53	2.10	.39	2.80	.40
Specific outlet	1.63	.61	2.04	.47	2.80	.40

Variable	Standardized betas		
	Husband	Joint	Wife
Wife employed	3.23	3.60	4.13
Size of carpet	2.23	1.46	1.44
How much to spend	.82	1.13	1.75
Visited stores	4.04	3.19	4.29
Specific outlet	2.21	2.45	2.92

Confusion matrix in percent				
Actual	Predicted			Total
	Husband	Joint	Wife	
Husband	60.4	35.4	4.2	48
Joint	17.4	70.1	12.5	92
Wife	1.8	10.9	87.3	55

## REFERENCES

- Cunningham, C.M. & Green, R.T. Purchasing roles in the U.S. family, 1955 and 1973. Journal of Marketing, 1974, 38, 61-64.
- Davis, H.L. Dimensions of marital roles in consumer decision making. Journal of Marketing Research, 1970, 7, 168-177.
- Davis, H.L. Measurement of husband-wife influence in consumer purchase decisions. Journal of Marketing Research, 1971, 8, 305-312.
- Davis, H.L. & Rigaux, B.P. Perception of marital roles in decision processes. Journal of Consumer Research, 1974, 1, 51-62.
- Engel, J.F., Kollat, D.T. & Blackwell, R.D. Consumer behavior. (2nd ed.) New York: Holt, Rinehart, & Winston, 1973.
- Ferber, R. Applications of behavioral theories to the study of family marketing behavior. Faculty working paper no. 125, University of Illinois, 1973.
- Ferber, R. & Lee, L.C. Husband-wife influence in family purchasing behavior. Journal of Consumer Research, 1974, 1, 43-50.
- Frank, R.E., Massy, W.F., & Morrison, D.G. Bias in multiple discriminant analysis. Journal of Marketing Research, 1965, 2, 250-258.
- Hansen, F. Consumer choice behavior: A cognitive theory. New York: The Free Press, 1972.
- Hempel, D.J. Family buying decisions: A cross cultural perspective. Journal of Marketing Research, 1974, 11, 295-302.
- Jaffe Associates. A pilot study of the roles of husbands and wives in purchasing decisions. New York: Life Magazine, 1965.
- Kenkel, W.F. Family interaction in decision-making on spending. In N.N. Foote (Ed.), Household decision-making. Volume 4. New York: New York University Press, 1961.
- Morrison, D.G. On the interpretation of discriminant analysis. Journal of Marketing Research, 1969, 6, 156-163.
- Sheth, J.N. A theory of family buying decisions. In J.N. Sheth (Ed.), Models of buyer behavior. New York: Harper and Row, 1974.
- Tigert, D. A psychographic profile of magazine audiences: An investigation of a media's climate. Unpublished paper presented to the Consumer Behavior Workshop. Columbus: Ohio State University, 1969.
- UPI (United Press International). Automakers say sales at 10-year October low. Columbia, South Carolina: The State, October 24, 1974.
- Woodside, A.G. Dominance and conflict in family purchasing decisions. Proceedings, Association for Consumer Research, 1972.





INTERGENERATIONAL PATTERNS OF CONSUMER BEHAVIOR<sup>1</sup>

Brent C. Miller<sup>2</sup>  
University of Minnesota

Questions are posed about how socialization in the parental family influences consumer behavior. Conceptual and methodological issues regarding age strata research in general are discussed, with particular emphasis being given to analyses between generations of the same family lines. Data are presented which show the line-age transmission of consumer patterns, especially a syndrome of poor consumership, from Hill's (1970) study of grandparents, parents, and their married children. Parallel intergenerational effects are suggested from a study in which consumer choices were simulated by parents and children playing a laboratory game, and from life-span developmental data where the economic context of the home was a major differentiating variable.

Many scholars of consumer behavior view the individual as the appropriate focus of analysis and pay little attention to the family as a decision-making unit or as an explanatory variable. The family has not been completely ignored by consumer researchers, but as Ferber (1973) noted, it has received relatively little attention. A possible exception to this oversight is the area of differentiation of marital roles in consumer behavior; such studies are fairly numerous. (See, for example, Sharp and Mott, 1956; Wolfstast, 1958; Blood and Wolfe, 1960; Davis, 1970 and 1971; Davis and Rigaus, 1974.)

One area which seems to have been left relatively unconsidered is that of intergenerational patterns of consumption. We might ask whether or not the following questions have been addressed by consumer researchers, and if so, how fully they have been answered: Are general patterns of spending, saving and management which are observed by children in the parental home reflected in similar behavior by these children after they become adults? To what extent are specific consumer choices made from generation to generation within family lines? Are there certain products or classes of products with high "brand loyalty" between generations? Apart from specific commodities, to what degree are global patterns of marital economic management and consumption -- styles of interacting, arbitrating, and deciding -- transmitted? To what degree are broad value orientations which impinge on consumer behavior (the conservative penny-pincher and impulsive whim satisfier) passed on between parents and children? Finally, relative to other types of explanatory variables, how influential in explaining consumership is family socialization? From these tantalizing substantive questions, let me turn our attention to some of the methodological issues involved in studying consumer behavior in different age strata.

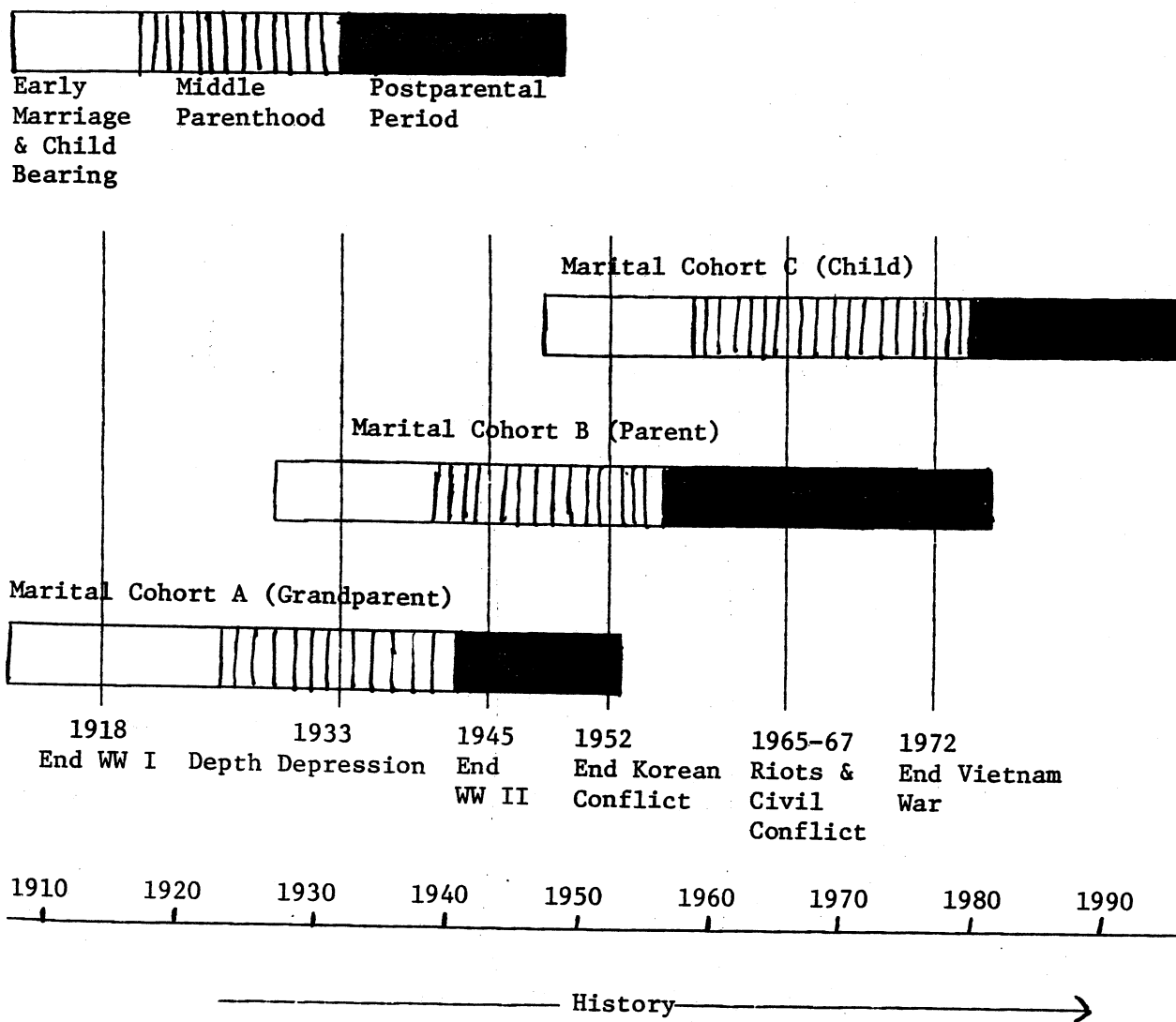
## Issues in Age Strata Research

A useful framework for discussing intergenerational issues is found in an excellent article by Bengtson and Black (1973). While there are many ways in which the term "generation" has been used by social scientists (see Troll, 1970), Bengtson and Black emphasize two meanings, and their dichotomy can be most helpful in analyzing intergenerational research issues related to consumer behavior. First, the term "generation" can be defined as a cohort; that is, an aggregate of individuals who share some common characteristic. In a now classic article on cohort analysis, Ryder (1965,845) defined cohort as "the aggregate of individuals (within some population definition) who experienced the same event within the same time interval." Most commonly then, cohorts are age-based, which means that the individuals within cohorts will have experienced together -- at about the same point in their personal developmental careers -- the impact of macrophenomena like wars, depressions, etc. Having similarly experienced such significant social events, the individuals in the cohort will thereby share some important and perhaps unrecognized characteristics that mark them as a social aggregate. Consumership analyses of changes and trends at the macro level would focus on cohort succession. For example, it appears that the generation of consumers who experienced the depression in their formative years has largely given way to a cohort of individuals who have known relative affluence for all of their lives. Cohort succession then is one basis for change in the social structure of the economic institution since patterns of money management, spending and saving are transformed by the turnover of generations. (The work by Eisenstadt, 1956, although not focused on consumer behavior, provides a nonempirical example of such macro analyses.)

At a second more micro level of analysis it might be observed that many consumer-related behaviors are transmitted between generations of the same family lines; that is, passed down from parent to child. Bengtson and Black have referred to this meaning of generation as lineage. Analyses cast at this micro level emphasize interpersonal interaction within family lines occurring in an individual or family developmental time frame, as opposed to cohort analyses in which there is a historical time dimension. In intergenerational analyses with the lineage focus, the emphasis is on individuals, the products of socialization, and the transmission of "private culture" within families.

To summarize and clarify these conceptions of generation, and to highlight the problems attendant in the study of age strata, Figure 1 has been included. This particular figure is from a paper by Hill (1974) patterned after a similar figure in Riley et al. (1972). Hill's figure appears here because it depicts three actual marital cohorts which will later be described in this paper. Each of the three cohorts represented is composed of couples married at approximately the same time. Riley (1973) points out that it is fundamental to understand that two basic processes occur simultaneously. First, within each cohort people age and individual marriages pass through various stages along their developmental courses. Second, new cohorts are continually formed as changes in the macro social structure occur as represented by the historical events along the time line. When family researchers take a cross sectional slice through such a layered age structure, the married couples in their sample differ in regard to both the stage of the family life cycle which they are at, and the marital cohorts to which they belong. As a consequence, differences which are observed between generations reflect a combination of marital career differences and cohort differences. Having briefly mentioned several key issues in the study of intergenerational phenomena, consumption-related data from several multi-generational analyses will now be presented. (For more in-depth

Stages in Family Development Cycle



Source: Hill (1973)

Figure 1. Selected Marital Cohorts Over the Developmental Cycle

treatments of these and related methodological issues, the reader is referred to excellent articles by Clausen, 1972; Riley, 1973; and Riley et al., 1972.)

Intergenerational Consumer Research

An impressive study of consumership across several generations is reported in detail by Hill and collaborators in Family Development in Three Generations (1970). Only a few summary statements can be presented here from the vast pool of intergenerational data available in the monograph. (See Hill, 1965, for a condensed account of this research.) Briefly, the sample consisted of 312 intact families linked through three generations living within the metropolitan area of the Twin Cities of Minneapolis and St. Paul. The data were obtained in four panel interviews scheduled over a twelve-month period covering the decisions undertaken in seven areas: residential location, remodeling, redeco-

rating, acquisition of automobile and durable goods, changes in financial investments and insurances and job changes.

In comparing the three generations it was no surprise to find educational and occupational upgrading from oldest to youngest generation. However, by comparing the percentage of each generation within lineages which were similar regarding various values, attitudes, and behavior, analyses were conducted to see to what extent there was intergenerational continuity or transmission. Religion was one heritage which appeared to be transmitted without marked change from generation to generation; the predominant pattern was stability within lineages, especially among wives of the married child generation where 80 per cent remained stable. In comparing broad value orientations between generations within the same family lines, the overall impression was one of similarity. To quote Hill (1970,44), "There appears to be high transmission of value orientations from generation to generation." Data on residential location suggested that "there is a strong tendency for neighborhood grades to be the same from generation to generation; that is, married child, parent and grandparent units will be living in similar type neighborhoods" (Hill, 1970: 31). By contrast, differences within family lineages were most apparent regarding conceptions about "proper" parent and child behavior. The shift to developmentalism was greatest from grandparents to parents, and persisted to the third generation. This pattern of greater differences between the two older generations is in line with the educational and occupational differences observed.

Looking more specifically at the decision process in consumership, Hill conceptualized the consumption cycle in four major phases: (1) the identification of unmet needs in the form of plans; (2) the process of choosing the best course of action; (3) an action of purchase or consumption; and (4) an evaluation of the action as satisfying or dissatisfying. The findings relative to phases one and three clearly showed how the generations differed in both the number of plans and actions taken during the year of study. (See Table 6.6 in Hill, 1965, and Chapter 9 in Hill, 1970.) The child generation was by far the most active, making more plans and taking more actions in the consumption areas enumerated, with the grandparent generation making fewest plans and taking fewest actions, and with the parent generation located between the oldest and youngest generations. However, plan fulfillment analyses revealed that the grandparent generation fulfilled the highest proportion of its less frequent plans. (See Table 6.8 in Hill, 1965.)

In phase two (the choice making process) a measure of rationality was devised based on the family's attempts to maximize information inputs before making a consumption choice. Examples of the criteria used were the extent of collection of information from family members, seeking alternatives, following a family policy, and seeking information from experts outside the family. The younger the generation, the higher the rationality tended to be. (See Chart 9.02 in Hill, 1970,205.)

At the evaluation phase satisfaction with outcomes was unexpectedly lower in the youngest generation, and yet this might have reflected a greater tendency to justify past actions among the older generations. Almost paradoxically, within each generation those who made the most plans, had the highest ratio of plan fulfillment, and preplanned most of their consumer actions, tended to be least satisfied with the outcomes. Only the rationality of the decision process turned out as expected to be highly predictive ( $r = .74$ ) of satisfaction with the consumption process and outcome. (See Chapter 10 in Hill, 1970.)

Perhaps the most salient evidence available for looking at the lineage transmission of patterns of consumership is Hill's data showing the number of plans which were fulfilled in combination with the number of actions taken which were preplanned. This composite consumership variable was used for a key analysis to determine the extent to which there was intergenerational continuity in consumer behavior. Operationally defined, similarity in the type of consumership for all three generations of the same family line constituted high continuity; similarity between any two generations of the same family line constituted moderate continuity, and no similarity among the generations was taken to indicate discontinuity.

Table 1 is presented to show the patterns of intergenerational continuity in the ratio of plans fulfilled and percentage of actions preplanned. The table reveals that low consumership (low plans fulfilled and low actions pre-

TABLE 1  
Consumer Types by Generational Continuity-Discontinuity:  
Three Generations Merged

Consumership Types	High	Moderate		Discontinuous	Total	
	(3 generation continuity) %	(2 generation continuity) %	(Deviant families) %	(All generations differed) %	f	%
Low Consumership						
Low plans fulfilled						
Low actions preplanned	60.0	32.5	42.5	30.3	72	35.8
Fruitless Planners						
Low plans fulfilled						
High actions preplanned	20.0	22.5	12.5	25.8	43	21.4
Unplanned Actors						
High plans fulfilled						
Low actions preplanned	—	2.5	30.0	19.7	27	13.4
High Consumership						
High plans fulfilled						
High actions preplanned	20.0	42.5	15.0	24.2	59	29.4
Total	100	100	100	100		100
N	15	80	40	66	201	

Source: Hill, 1970, 233

planned) is more likely to persist over three generations than any other behavior pattern. Nine of the fifteen families in three generation lines who had similar consumership behavior were of this indiscriminate planning, impulsive acting-without-planning type. High consumership is transmitted, however, since 34 of 59 (62.5 per cent) high consumership families were alike over two or more generations. These data present the most convincing evidence available that important consumer behaviors, particularly at the impulsive end of the scale, are transmitted intergenerationally.

A related bit of evidence can be drawn from a game-simulation called SIMCAR (standing for SIMulation of CAREer patterns) which has been developed at the University of Minnesota for studying family interaction under laboratory conditions (Tallman et al., 1974). It is essentially a simplified model of career and consumption choices which typically are made by and for children in late adolescence and early adulthood. SIMCAR is constructed on the basis of an economic consumption model so that the most points are earned by players if they choose alternatives which maximize financial return. Indeed, points earned are redeemable for real money at the end of the game. Doing well in the game implicitly emphasizes deferred gratification and occupational-career orientations. Players are given play money and tokens representing time, and they are allowed to invest as much of each as they wish in consumption, religious, marital, educational, and occupational choices.

There are two halves in the simulation-game; in the first half parents play the game with the child, and in the second half the child plays the game alone. The experimental manipulation is the creation of a success or failure condition in the family game by means of unplanned event cards which are interjected into the game as a normal feature. What the families do not know is that "success" or "failure" conditions are created by stacking the cards for or against them in such a way that the unplanned event cards either impede or facilitate their success at the game. Consequently, the child's game which is played alone can be influenced by the family failure or success in the previous game.

One fascinating finding is that boys (in Minneapolis and from villages and a city in Mexico) who experienced the failure conditions with their families scored much lower in the second round than those who had experienced the success condition (Tallman and Wilson, 1974). It would be even more impressive if the failure and success conditions produced lower and higher scores respectively than the control or neutral condition. Nevertheless, the findings suggest that, over the short term at least, boys who experienced "failure" with their parents while engaged in a simulation of adolescent consumption and career choices were influenced to follow a less economically rational route when going through the simulation a second time alone. If this effect could be shown to be long lasting and persistent, it would seem to imply that real-life consumption experiences could be influenced by earlier "failures" in the family context. Of course this is the implication of Hill's data showing intergenerational transmission of low consumership.

Elder's (1974) fascinating cohort analysis of the Children of the Great Depression is a repository of rich life course data related to these issues. Members of the panel were born in 1920-21, and with the most recent assessments in the mid-1960's, Elder has constructed a splendid account of the impact of the depression experience on the life courses of these individuals. Trying to extract only a single overall generalization or so is inadequate,

but a primary finding seems to be that for both men and women raised in depression-deprived homes, the importance of the family and children is significantly higher than for those from less deprived homes.

We can answer affirmatively to only a few questions posed at the beginning of this paper. There is evidence of intergenerational transmission of various consumer behaviors. Particularly striking is the continuity between generations in syndromes of unfulfilled plans and/or making consumer actions impulsively. However, there are also a great many cohort and developmental subtleties which make analyses of lineage continuities very complex.

One of the neglected issues is that intergenerational studies which take marriages or families as the unit of analysis really should consider prior generational patterns coming out of two different kin structures -- those of both the husband and wife. Each of the marital partners brings with them different family backgrounds and socialization experiences (Hill and Aldous, 1969). This fact probably reduces considerably the lineage transmission of consumer patterns. Since divergent patterns will usually be negotiated and compromised by marital partners, the degree of difference in the lineage backgrounds is likely to significantly affect continuities between vertical generations.

Finally, while it is sensible to consider parental socialization of children into behavior patterns, we should not forget that intergenerational transmission occurs bilaterally -- socialization is reciprocal. How much are parental and marital consumption patterns altered by the presence and characteristics of children? What young adult has not attempted (sometimes successfully) to get his parents to try some new product? It seems inescapable that the further study of the complexities of intergenerational relations will produce answers to interesting and important questions, and provide a more complete picture of why families and consumers behave as they do.

#### FOOTNOTES

<sup>1</sup>Appreciation is expressed to Reuben Hill, Robert Ferber and David Klein for reading and commenting on an earlier version of this paper.

<sup>2</sup>The author is a graduate student in the Department of Sociology and affiliated with the Family Study Center at the University of Minnesota, 1014 Social Science Building, Minneapolis, Minnesota 55455.

#### REFERENCES

- Bengtson, V. L. & Black, D. K. Intergenerational relations and continuities in socialization. In P. Baltes and W. Schaie (Eds.), Life-span developmental psychology: Personality and socialization. New York: Academic Press, 1973.
- Blood, R. O. & Wolfe, D. M. Husbands and wives: The dynamics of married living. New York: Free Press, 1960.
- Clausen, J. The life course of individuals. In M. Riley, M. Johnson & A. Foner (Eds.), Aging and society III: A sociology of age stratification. New York: Russell Sage, 1972.



- Davis, H. L. Dimensions of marital roles in consumer decision making. Journal of Marketing Research, 1970, 7, 168-177.
- Davis, H. L. Measurement of husband-wife influence in consumer purchase decisions. Journal of Marketing Research, 1971, 8, 305-312.
- Davis, H. L. & Rigaus, B. P. Perception of marital roles in decision processes. Journal of Consumer Research, 1974, 1, 51-62.
- Eisenstadt, S. N. From generation to generation. Glencoe, Illinois: Free Press, 1956.
- Elder, G. H., Jr. Children of the great depression: Social change in life experience. Chicago: University of Chicago Press, 1974.
- Ferber, R. Family decision making and economic behavior. In E. B. Sheldon (Ed.), Family economic behavior. Philadelphia: J. B. Lippincott, 1973.
- Hill, R. L. Decision making and the family life cycle. In E. Shanas & G. Streib (Eds.), Social structure and the family: Generational relations. Englewood Cliffs, New Jersey: Prentice-Hall, 1965.
- Hill, R. L. Family development in three generations. Cambridge, Massachusetts: Schenkman, 1970.
- Hill, R. L. Social theory and family development. In J. Cuisenier (Ed.), Le cycle de la vie familiale dans les sociétés européennes. Paris: Mouton, 1974.
- Hill, R. L. & Aldous, J. Socialization for marriage and parenthood. In D. Goslin (Ed.), Handbook of socialization theory and research. Chicago: Rand McNally, 1969.
- Riley, M. W. Aging and cohort succession: Interpretations and misinterpretations. Public Opinion Quarterly, 1973, 37, 35-49.
- Riley, M. W., Johnson, M., & Foner, A. Elements in a model of age stratification. In M. W. Riley, M. Johnson, & A. Foner (Eds.), Aging and society III: A sociology of age stratification. New York: Russell Sage, 1972.
- Ryder, N. The cohort as a concept in the study of social change. American Sociological Review, 1965, 30, 843-861.
- Sharp, H. & Mott, P. Consumer decisions in the metropolitan family. Journal of Marketing, 1956, 21, 151-158.
- Tallman, I. & Wilson, L. Simulating social structures: The use of a simulation game in cross-national research. Simulation and Games, 1974, 5, 147-167.
- Tallman, I., Wilson, L. & Straus, M. SIMCAR: A game simulation method for cross-national family research. Social Science Information, 1974, 13(2), 121-144.

Troll, L. The family of later life: A decade review. In C. B. Broderick (Ed.), A decade of family research and action. Minneapolis: National Council on Family Relations, 1970.

Wolfgast, E. H. Do husbands or wives make the purchasing decisions? Journal of Marketing, 1958, 23, 151-158.



METHODOLOGICAL CONSIDERATIONS IN FAMILY DECISION-MAKING STUDIES

Marilyn M. Dunsing and Jeanne L. Hafstrom<sup>1</sup>  
University of Illinois

Studies of decision-making present numerous problems. The methodological problems encountered in data collection, in choice of participants, and in decision-making concept are examined. The idea presented is that dissimilarity in results among studies may be explained: 1) in terms of what method is used to collect data--the survey or self-report versus the behavior or observational method; 2) in terms of who is interviewed--husband, wife, children or some combination, thereof, as well as who does the interviewing--male or female; and 3) in terms of the concept itself--hypothetical versus real problems, what decisions are studied, and decision-making stages. Suggestions are made concerning additional research needed if the methodological problems involved in decision-making are to be solved.

Increasing recognition has been given to the methodological problems encountered in decision-making studies investigating "influence" or "power" in families (Hallenbeck, 1966; Olson, 1969; Safilios-Rothschild, 1969 and 1970; Davis, 1970; Turk & Bell, 1972; Davis & Rigaux, 1974). The purpose of the present paper is to examine the methodological problems encountered; first, in data collection; second, in choice of participants; and third, in the decision-making concept itself.

#### Data Collection

Methodological problems in data collection exist for both the survey or self-report method and the behavioral or observational method. Specific problems include those of cost, types of power, nature of the decision-making concept, measurement of the concept, and stages of the concept.

The survey method is much less costly than the observational method. This means that with a fixed amount of money a much larger sample can be obtained using the former method. However, cost should not be the deciding factor unless the two methods give equally good results. Then, of course, the less costly method would be used, all other things being equal. But all other things are not equal. "Other" things depend on the decision-making model being used.

How is decision-making measured? Who are the respondents? What are the independent variables? Which method is better to use depends on the answers to these questions. For example, if the model calls for the respondent's perception of who usually wins when there is a disagreement as the measure of decision-making, if the model calls for only the wife being the respondent, and if the model calls for the use of "traditional" economic-demographic variables to explain influence in decision-making, then clearly the survey method is the one to use.

On the other hand, if the model calls for a consensus on the course of

action taken when a disagreement exists between respondents as the measure of decision-making, and for interaction variables to be used in addition to economic-demographic variables in explaining influence in decision-making, then clearly the behavioral method is the one to use. It should be apparent that if the model calls for the use of the behavioral method and if funds are insufficient for this method, switching to the survey method would not be appropriate until a reassessment of the model takes place.

Researchers critical of the survey method contend that predicted power, which is based on the respondent's perception of the situation, is not closely related to actual or real power. This contention needs to be evaluated. We know that people are influenced by their perceptions. For example, in a recent study dealing with satisfaction with level of living (Hafstrom & Dunsing, 1973) the findings indicated that the homemaker's perception of the adequacy of family income was a more important determinant of satisfaction than actual income. Since in this study perception was significantly related to satisfaction, why should it not also be related to action?

If the wife perceives erroneously that she has the authority to make a particular decision, such as the purchase of a dishwasher, and executes this decision, then from the point of view of market behavior her perception has become reality. If the husband does not object to her purchase, then family "power" with respect to this item may have changed, or her husband's perception of the situation may be that joint decision-making has taken place.

Those researchers critical of the survey method also contend that since it contains only the respondent's perception of who makes the decision, the observational method should be used to obtain the actual or real power. These critics need to recognize that not only are the respondents reacting according to their perceptions but also observers (if used) are making judgments about the respondents' behavior in terms of their own perception of the situation. The relevant question appears to be whose perceptions--respondents' or observers'--come closer to reality?

Those critical of the survey method argue that people are more likely to be influenced by socially-desired norms when they respond to questionnaires (in front of an interviewer) than when they are reacting with others to come to a consensus about a disagreement. The validity of this argument is questionable if the interaction among family members is taking place in front of an observer. Under these conditions, there is no a priori reason to assume that the behavior of respondents would be influenced by conventional norms any differently in the behavioral than in the survey situation. If behavior is in fact influenced, it appears more likely that influence would occur in the behavioral rather than in the survey situation. This is because interaction takes place in public in the behavioral situation.

A number of studies have compared the results based on self-reporting with those based on the observational method. In general, the results do not show much agreement (Kenkel, 1963; Olson, 1969; Turk & Bell, 1972). Why is this the case? Is it because one method is in fact superior to another method, or is it because the results are based on different concepts of decision-making, and, consequently, should not be compared? In other words, if two studies are compared, dissimilarity in results may be explained in terms of differences in decision-making concepts rather than in terms of differences in method used. We will return to this problem later.

Dissimilarity in results between two studies also may be explained in terms

of differences in measurement of the same decision-making concept. For example, the survey method uses the question "Who mainly decides--husband, wife, or both--when a disagreement occurs in purchasing a refrigerator?" What measurement of the concept is analogous using the behavioral method? Should the concept be measured by the amount of talking done by each participant? Should it be measured by who first suggested buying the article or by who made more suggestions regarding the purchase? Should it be measured by comparing the husband's and wife's instrumental acts, by comparing their expressive acts, or by comparing the husband's instrumental acts with the wife's expressive acts? We will return to this problem later.

Perhaps studies that have used different methods had dissimilar findings because different stages of the concept were measured. For example, the survey method may have investigated who dominated in recognizing that a problem existed in savings objectives, while the observational method may have investigated who dominated in making the final decision. Results may differ mostly because of noncomparability of stages in the family decision-making process. Again, we will return to this problem later.

These examples indicate that whether the behavioral method or the survey method is preferable depends on a number of different factors, many of which have been examined. When studies, which have used different methods, have been compared in the past it appears likely that differences in addition to method, such as meaning or measurement of the decision-making concept, have also existed. These differences make a comparison between the methods very difficult.

It has been suggested that which method to use--self-reporting or observational--should depend on the dimensions of decision-making that are being studied (Safilios-Rothschild, 1970). If this is the case, research is needed to determine which dimensions are best measured by self-reporting techniques and which are best measured by observational techniques.

#### Choice of Participants

Regardless of whether the survey or observational method is used, a major problem is the choice of participants. Does it matter who is interviewed or observed when investigating "influence" or "power" in families? Should it be the wife, or husband, or both? Should it be only the children or just one of them? Should it be both the children and their parents? The majority of self-reporting studies in the past have used the wife, primarily because of her availability and, therefore, reduced cost per schedule. However, comparisons of responses made by husbands and by wives indicate differing results.

When the responses of all husbands are compared with those of all wives (mean scores), the findings are similar for many studies (Wilkening & Morrison, 1963; Davis, 1970; Granbois & Willett, 1970). However, when the responses of wives are compared with those of their own husbands, the results vary considerably depending on what decisions are being investigated. For example, in a recent study (Davis & Rigaux, 1974, p. 58) while 82 percent of the husbands and wives agreed on who made the final decision about housing (in terms of location, purchase price, or rent), only 49 percent agreed on who made the final decision about other furnishings (rugs, draperies).

Inasmuch as the responses of husbands and their wives differ, what is the nature of the bias introduced by the use of one or the other? Some studies (such as Ferber, 1955; Heer, 1962) have found that individuals tend to

attribute less influence to themselves than is attributed to them by others. Other studies (such as Heer, 1962; Safilios-Rothschild, 1969) have found that while the husband was more likely to attribute "influence" to his wife than she was, the wife was more likely to perceive that they had equal "influence." In yet another study (Granbois & Willett, 1970), couples who disagreed (about half) were about equally divided between those where the wife's estimate of "influence" exceeded that of her husband's and those where the husband's estimate exceeded that of his wife's. These conflicting findings point to the need for additional research on the consequences of husband-wife participation in decision-making studies.

In some studies using the self-reporting method only children were used. Again, this is primarily because of availability and cost considerations. The feeling is that the results could be misleading if only children's responses are relied upon to obtain information about who makes the decisions in the family. Some evidence indicates that the characteristics of the children will influence the results. For example, Hess and Torney (1962) found that younger and older children were likely to perceive the family "influence" structure differently. In addition, they (as did Bowerman & Elder, 1964; and King, 1969) found that sex influenced the children's responses. Boys tended to perceive more father "influence" than girls, whereas girls tended to perceive more mother "influence" than boys. These findings indicate the need for more research concerning the effect of children's participation in decision-making studies investigating "influence" in families.

Another participant in the survey or observational method is the interviewer or observer. Much has been written about interviewer bias. It is recognized that bias can occur in terms of asking questions, interpreting questions, recording respondents' answers to open-ended questions, and giving "cues" to respondents, particularly when open-ended questions are used. Proper interviewing training should do much to eliminate this type of bias. However, another type may exist. Does the sex of the interviewer or observer influence the results?

One study (Kenkel, 1961) found some evidence that the sex of the observer did influence the respondents, particularly their interactive behavior. In this study, when a female observer was used the wives tended to have more influence and to perform more in the areas of giving ideas and suggestions than when a male observer was used. In another study (Feldman & Rand, 1965), the sex of the interviewer did not influence which sex "won" the decision. Additional studies are needed to determine whether the sex of the interviewer or observer does or does not influence the results. If sex does influence the results, then additional studies are needed to distinguish the effect of the interviewer's or observer's sex from the effects of the method used and the decisions studied.

#### Decision-making Concept

Methodological problems related to the concept of decision-making are many and varied. One question often raised in the past is whether decision-making is unidimensional or multidimensional in nature. Another question is can families be characterized as husband-dominated, wife-dominated, or equalitarian. One investigation (Farber, 1949) divided decisions into policy and routine. Others (such as Kenkel, 1957; Turk & Bell, 1972) used instrumental and expressive acts as measures of "power." As a result of findings from numerous studies (Safilios-Rothschild, 1969; Davis, 1970; Turk & Bell, 1972), the multidimensional nature of decision-making has become apparent.

In studies using the survey method, the specific question or questions asked as a basis for measurement of decision-making have varied considerably. For example, in one study respondents were asked who usually won when there was a disagreement between the spouses (Heer, 1958), while in another study respondents were asked who is the real boss in your family (Hess & Torney, 1962). Turk and Bell (1972, p. 221) obtained responses to both questions from the same group of husbands and wives. The correlation between these two measures was .40.

Turk and Bell (1972) also obtained responses to a group of decision questions used by Blood and Wolfe (1960). The questions were concerned with jobs, cars, vacations, life insurance, wife working, doctors, dwelling unit, and food. Turk and Bell combined the answers to these questions, as Blood and Wolfe had done, to arrive at one over-all score. The correlation between the Blood and Wolfe measure of decision-making and the question who usually won was only .17, but between the Blood and Wolfe measure and the question who is the real boss it was .54 (Turk & Bell, 1972, p. 221). These results raise doubts about whether the three sets of scores were measuring the same dimension of decision-making. For example, does the question who is the real boss in your family indicate influence or authority?

Doubts have also been raised about combining answers to numerous questions in order to obtain a "global" or "over-all" score to measure dimensions in decision-making. In the case of Blood and Wolfe all questions were combined, which indicated that the researchers assumed that they all measured the same dimension of decision-making. In addition, Blood and Wolfe gave all statements the same weight. Is either of the assumptions appropriate? Do all statements measure the same dimension? Do all statements represent areas of equal importance to the family? Safilios-Rothschild (1969) suggests that factor analysis be used on a series of statements to determine which of a number of specific decisions to include in a study of decision-making.

In studies using the behavioral method, variation has also occurred in the situations used as a basis for measurement of decision-making. In one study (Kenkel, 1961), husbands and wives were asked to come to a consensus on how they would spend a hypothetical gift of \$300. In another study (Olson & Ryder, 1970), husbands and wives were asked to come to a consensus on 18 vignettes which presented various types of marital conflicts (such as wife's purchase of new shoes, husband's forgetfulness about throwing out the garbage, decision to have a baby). While hypothetical situations were used in both studies, the situation applied to the family itself in the former study but to other families in the latter study.

In yet another study (Olson, 1969), couples (who were expecting their first child in a few months) were asked to come to a consensus when differences existed on those items referring to decisions which might be encountered by couples the first few weeks after the birth of their first child. In another study (Hollenbeck, 1972), husbands and wives were asked to come to a consensus when differences existed about the timing of purchases within the next three years on items costing \$50 or more which had been suggested individually and independently by them.

Numerous questions are raised by the examples presented above. Of what significance are the results when hypothetical rather than real problems of the participants are used? Of what significance are the results when the problems used pertain to other families rather than to the families participating in the study? Are the situations presented to the participants such



that they become sufficiently involved in them so that they will play their typical roles? Do the situations represent areas of equal frequency in enactment?

The commonly used measure of dominance in decision-making in behavioral studies is Strodbeck's (1951) revealed preference technique (RPT) or a modification of it. In the RPT each participant fills out an individual form, the responses of the participants are compared, participants are asked to come to a consensus on the items, and "winning" is measured by the proportion of instances in which the responses of each participant prevail. Both the Kenkel (1961) and the Olson and Ryder (1970) studies use "winning" as their measure.

No matter how dominance in decision-making is judged, the importance of the decision for each participant should be determined or any conclusions reached by the researcher may be erroneous. Of what importance is it to decision-making if the wife wins because her husband couldn't care less about the outcome? In Hollenbeck's study, influence was measured by finding out not only the proportion of instances in which the opinion of each participant prevailed but also the relative strength of feeling of each participant about this decision.

Irrespective of whether the survey or behavioral method is used, the researcher is faced with the problem of deciding on the decision area or product category to be studied. To illustrate, the decision area could be very broad such as material and non-material components. Or the decision area could involve different types of non-material components such as child rearing and marital relations, or different types of material components such as spending and saving. The decision area could be confined to different kinds of saving such as savings accounts, government bonds, and stock, or to broad categories of spending such as durable goods, nondurable goods, and services. The decision area could be narrowed still farther to specific types of durable goods such as automobiles, washers, dryers, and refrigerators.

If the wife is found to dominate in the area of child rearing, can she also be expected to dominate in the area of marital relations? If the husband is found to dominate in the purchase of an automobile, can he also be expected to dominate in the purchase of a refrigerator? If he is found to dominate in the decision to save, can he also be expected to dominate in the decision to open a savings account? Differences in decision areas used may explain, wholly or in part, conflicting findings between studies concerning dominance in decision-making.

Another problem concerns stages in the decision-making process. In the past, researchers have not tended to recognize that the decision-making process contains a number of stages. In addition to specifying which decision areas are to be studied, researchers also need to specify the stage or stages of each decision area to be studied. A recent study (Davis & Rigaux, 1974) obtained decision information for three stages--problem recognition, search for information, and final decision--for 25 decisions varying from specific consumer goods to the child's school and program of study. Their findings indicate that for certain types of decisions the results vary considerably depending on the stage used. For example, with respect to automobiles, in the search for information stage the decision was husband-dominant, whereas in the final decision stage it was syncretic (more than 50 percent of the families made the decision jointly). It would appear that more research is needed to identify relevant stages in the decision-making process for numerous types of decisions.

Still another problem concerns the independent variables used. Economic-demographic variables have tended to be emphasized in the past. Perhaps more imagination is called for in the use of different types of independent variables in future studies. In the Hollenbeck study (1972, pp. 71-72), the model included interaction variables (Bales) and personality variables (Cattell), as well as economic-demographic variables. In terms of beta coefficients, seven of the personality variables were significant at or beyond the .05 level compared to two of the interaction variables and one economic-demographic variable.

Husband's total hours away from home, the significant economic-demographic variable, is not generally used in decision-making studies. More traditional variables, such as income (husband's, wife's, total), husband's occupation, wife's employment status, and education and age of husband and wife, were used in the Hollenbeck study but were not significant. Perhaps consideration should be given to economic-demographic variables other than those generally used in future decision-making studies. While interaction variables call for the use of the behavioral method, personality variables do not. Hollenbeck's findings suggest that more consideration be given to the use of personality variables in future studies regardless of the method used.

In the past, many studies have interpreted decision-making findings as indicating power in the family. Safilios-Rothschild (1970) feels that this interpretation is not warranted. She feels that decision-making is only one measure of power; authority and influence are two additional measures. If this is the case, what are the implications for decision-making research?

In summary, it would appear that studies of decision-making present numerous problems. These include such things as selecting the method to use, deciding on the dimensions to study in terms of products and stages, deciding on who will be interviewed, and deciding on what independent variables to use. Much research effort is needed in order to solve these problems. Hopefully, as a result of future research in the area of family decision-making, it will be possible to make generalizations about decisions, particularly in the consumer area.

#### FOOTNOTE

1. Marilyn M. Dunsing is Professor of Family and Consumer Economics and Jeanne L. Hafstrom is Assistant Professor of Family and Consumer Economics, School of Human Resources and Family Studies, University of Illinois, Urbana-Champaign Campus, Urbana.

#### REFERENCES

- Blood, R. O., Jr., & Wolfe, D. M. Husbands and wives, the dynamics of married living. Glencoe, Ill.: Free Press of Glencoe, 1960.
- Bowerman, C. E., & Elder, G. H., Jr. Adolescent perception of family power structure. American Sociological Review, 1964, 29, 551-67.
- Davis, H. L. Dimensions of marital roles in consumer decision making. Journal of Marketing Research, 1970, 7, 168-77.
- Davis, H. L., & Rigaux, B. P. Perception of marital roles in decision

- processes. The Journal of Consumer Research, 1974, 1 (2), 51-62.
- Farber, B. A study of dependence and decision-making in marriage. Unpublished doctoral dissertation, University of Chicago, 1949.
- Feldman, H., & Rand, M. Egocentrism-altercentrism in the husband-wife relationship. Journal of Marriage and the Family, 1965, 27, 386-91.
- Ferber, R. On the reliability of purchase influence studies. Journal of Marketing, 1955, 19, 225-32.
- Granbois, D. H., & Willett, R. P. Equivalence of family role measures based on husband and wife data. Journal of Marriage and the Family, 1970, 32, 68-72.
- Hafstrom, J. L., & Dunsing, M. M. Level of living: Factors influencing the homemaker's satisfaction. Home Economics Research Journal, 1973, 2, 119-32.
- Heer, D. M. Dominance and the working wife. Social Forces, 1958, 36, 341-47.
- Heer, D. M. Husband and wife perceptions of family power structure. Marriage and Family Living, 1962, 24, 65-67.
- Hess, R. D., & Torney, J. V. Religion, age and sex in children's perceptions of family authority. Child Development, 1962, 33, 781-89.
- Hallenbeck, P. N. An analysis of power dynamics in marriage. Journal of Marriage and the Family, 1966, 28, 200-03.
- Hollenbeck, C. L. Husband-wife influence in purchasing decisions. (Doctoral dissertation, University of Illinois), Ann Arbor, Mich.: University Microfilms, 1972. No. 72-19, 848.
- Kenkel, W. F. Influence differentiation in family decision making. Sociology and Social Research, 1957, 42, 18-25.
- Kenkel, W. F. Family interaction in decision-making on spending. In N. N. Foote (Ed.), Household decision making. New York: N.Y.U. Press, 1961. Pp. 140-164.
- Kenkel, W. F. Observational studies of husband-wife interaction in family decision-making. In M. B. Sussman (Ed.), Sourcebook in marriage and the family. (2nd ed.) Boston, Mass.: Houghton-Mifflin, 1963. Pp. 144-156.
- King, K. Adolescent perception of power structure in the Negro family. Journal of Marriage and the Family, 1969, 31, 751-55.
- Olson, D. H. The measurement of family power by self-report and behavioral methods. Journal of Marriage and the Family, 1969, 31, 545-50.
- Olson, D. H., & Ryder, R. G. Inventory of marital conflicts (IMC): An experimental interaction procedure. Journal of Marriage and the Family, 1970, 32, 443-48.
- Safilios-Rothschild, C. Family sociology or wives' family sociology? A cross-cultural examination of decision-making. Journal of Marriage and

- the Family, 1969, 31, 290-301.
- Safilios-Rothschild, C. The study of family power structure: A review 1960-1969. Journal of Marriage and the Family, 1970, 32, 539-52.
- Strodtbeck, F. L. Husband-wife interaction over revealed differences. American Sociological Review, 1951, 16, 468-73.
- Turk, J. L., & Bell, N. W. Measuring power in families. Journal of Marriage and the Family, 1972, 34, 215-222.
- Wilkening, E. A., & Morrison, D. E. A comparison of husband and wife responses concerning who makes farm and home decisions. Journal of Marriage and Family Living, 1963, 25, 349-51.



SOME UNANSWERED QUESTIONS ON FAMILY DECISION-MAKING<sup>1</sup>

Robert Ferber<sup>2</sup>  
University of Illinois

This summary paper outlines directions for future research on family decision-making, based on the work to date.

To supplement the other papers presented at this session, this note seeks to highlight some unanswered questions on family decision-making which would seem to deserve further study. These questions are organized under the four major topics covered in this workshop, namely, the nature of decision-making, its determinants, its effects, and the content and methodology of decision-making studies. The questions listed in this note are not meant to be exhaustive but rather to indicate those aspects of the subject that seem especially worthy of future study.

#### I. The Nature of Decision-Making

The work that has been done on the descriptive aspects of family decision-making raises about as many questions as it answers. Among these questions are the following:

1. To what extent does the family act as a unit in making various types of decisions? In economic theory, the family is treated implicitly as a unit, though evidence is accumulating that this assumption is not really valid. Indeed, some of the models on economic behavior, such as on labor supply, are beginning to assume that the different members of the family act as individuals with their own goals and objectives, and that these goals are more important for work behavior than any attitudes or preferences of the family as a group.<sup>3</sup>
2. What is the role of family members other than the husband and the wife in decision-making? Some work on this question has been done on the influence of children, and of teenagers especially.<sup>4</sup> The influence of other family members seems to have been virtually ignored. Nevertheless, in one study some time ago, Elizabeth Wolgast found that in 5% of the families, members other than the spouse were responsible for keeping track of money and of paying bills.<sup>5</sup> Considering that many families do not have more than two members, this proportion must very likely be far higher if an adjustment for this fact were made.
3. To what extent are decisions interrelated within a family? Thus, is the decision maker on one type of decision also likely to be the decision maker on other decisions and, if so, under what circumstances? Harry Davis has provided much evidence that as a rule any particular economic or marketing purchase decision is actually composed of a sequence of decisions, and that the key person at one stage of a sequence may not be the same as the key person at another

stage.<sup>6</sup> To what extent this is so among different types of decisions altogether has hardly been touched upon.

4. While purchase decisions have been explored fairly extensively, as is evident from the foregoing references, other areas of family decision-making seem to have been largely ignored. This includes decision-making with regard to such aspects as handling savings and investments, career choice, and even fertility.
5. What is the role of decision-making on impulse behavior? The question itself may seem like a contradiction, but the studies that have investigated the circumstances surrounding so-called impulse purchases suggest that many of these purchases are not so unplanned as would seem from initial observation, in the sense that prior discussion regarding the possibility of a purchase may have already taken place but the purchase itself was triggered by an event unexpected at that particular time.<sup>7</sup> In effect, within a decision-making framework, "impulse" behavior may be less impulsive or irrational than appears at first sight.

## II. Determinants of Decision-Making

1. A fair amount of work has been done in recent studies on the relative importance of different factors that seem to influence who is the decision maker.<sup>8</sup> Still, much more work needs to be done on the influence of different characteristics of the individuals in the family, and the family structure, on the identity of the decision maker. Thus, all indications are that the prevalence of joint decision-making by husband and wife is on the increase, and especially so for the younger families. From a macro point of view, there are many reasons why this is to be expected. From a micro point of view, however, we do not seem to have much evidence on what factors within the family are bringing about this increase, and in particular on the types of families which are more likely to be characterized by joint decision-making.
2. A related question is, what determines how the identity of the decision maker is changed? Obviously, habit plays a major role. Obviously also, some change has to take place either in the structure of the family or in the outcome of decision processes for such a shift to take place. The fact remains, however, that this question does not seem to have been studied, and we have little indication of how these changes take place or under what circumstances. For example, are they brought about by changes within the family or by external events, and how?
3. To what extent are intergenerational effects present in decision-making? Brent Miller has reported, based on the pioneering study by Reuben Hill and his associates, that intergenerational effects do seem to be present.<sup>9</sup> However, this study is primarily descriptive and does not attempt to take into account factors relating to the characteristics and the structure of the present family. If these factors were incorporated within a multivariate framework, would an intergenerational effect still exist?

4. What is the influence of friends and relatives not in the household on family decision-making? That these influences can be substantial was brought out many years ago in the study by Katz and Lazarsfeld,<sup>10</sup> but relatively little work has been done on this since that time.
5. A related question is the influence of search for information on decision-making. While much work has been done on this subject,<sup>11</sup> it is still not clear whether this search serves mainly to support a preconceived idea or to alter the plans of a family; and under what circumstances is each likely to be true.
6. As was pointed out in the workshop, one population group that has received very little attention from the point of view of decision-making is the elderly. All of the emphasis seems to be on decision-making by the young or by husband-wife families in relatively early stages of the life cycle. It would be very useful to have corresponding studies on decision-making by the elderly, both by couples and by single individuals, including both the nature as well as the determinants of decision processes by these groups.

### III. Effects of Decision-Making

1. Very little attention as yet has been given to what in many ways may be termed as the key question underlying the study of family decision-making. This is, what difference does it make on consumer behavior whether one member of a family or another is the decision maker in a particular case? To what extent, if any, does the behavior of a family depend on the identity of the decision maker? In a study by Lucy Lee and myself that looked into this aspect, certain types of purchase behavior were indeed found to differ depending on who made the final decision, even after other relevant variables were taken into account.<sup>12</sup> However, this was only one study on a particular type of population, and this key question remains as yet largely unexplored.
2. What is cause and what is effect in family decision-making? The nature of this question is perhaps best illustrated with reference to a recent study at the University of Illinois that found one of the principal determinants of the wife's working to be the attitude of the husband: if the husband had a favorable attitude toward his wife working, she was much more likely to be so engaged.<sup>13</sup> What could not be established from this study is, which came first? In other words, were the wives working as a result of the favorable attitudes toward this activity of their husbands? Or, had the wives "selected" their husbands using this attitude as a criterion, so that in effect the relationship is the other way around!

### IV. Methodology of Decision-Making

1. While some attention has been given to means of collecting data on decision-making, most of the emphasis in past studies has been heavily on the substantive results. As a result, questions such as the effect of the presence of an interviewer on information supplied



on decision-making has not been explored. For example, is more reliable information obtained by a self-administered approach without any outside person present, by perhaps separate and simultaneous interviews with different family members, or by some other method?

2. What types of family members and under what conditions are they accurate reporters of the influence of themselves and of other family members in decision-making? As Marilyn Dunsing has reported, while results on family influence seem to be similar in the aggregate regardless who is interviewed, substantial discrepancies are obtained on a micro basis, that is, for particular families for certain types of consumer actions. Hence, a corollary question to the previous point becomes how one can determine in advance whether a particular method of data collection will be biased and, if so, in what direction.
3. Is the recall approach the best way of obtaining information? While this is the most commonly employed method in decision-making studies, and is in many ways the easiest, no comparative study seems to have been made of different approaches on the same sample or on matched samples. From a data collection point of view, the door seems wide open for imaginative work on this question.

These are by no means all of the questions that could be raised on the subject, and others have been suggested by the foregoing papers. Taken together, it is hoped that this workshop not only helped provide further information on family decision-making but has also pointed the way to avenues that are particularly worthy of further investigation.

#### FOOTNOTES

1. The thoughts expressed in this paper are based in part on work undertaken under a grant from the National Science Foundation, Grant GS-36166, for which the author would like to express his appreciation.
2. Research Professor of Economics and Business Administration, and Director of Survey Research Laboratory, University of Illinois.
3. See, for example, the model in Ashenfelter and Heckman (1974).
4. See, for example, Cateora (1961) and also Gibbs (1963).
5. See Wolgast (1958).
6. See Davis (1970); also Jaffe and Senft (1966).
7. See Granbois (1971). Other unpublished studies have had similar findings.
8. For example, Hempel (1974); also Ferber and Lee (1974).
9. See Hill (1970) for such a study in the Minneapolis area.
10. See Katz and Lazarsfeld (1955), which covered a wide variety of consumer decisions.

11. For example, Cox (1967).
12. See Ferber and Lee (1974); the focus here was mainly on financial decisions.
13. See Sampson (1972). The sample, however, was restricted to the Champaign-Urbana area.

## REFERENCES

- Ashenfelter, O. Heckman, J. The estimation of income and substitution effects in a model of family labor supply. Econometrica, 1974, 42, 72-86.
- Cateora, R. R. An analysis of the teenage market. Austin: University of Texas Bureau of Business Research, 1961.
- Cox, D.F. (Ed.) Risk-taking and information handling in consumer behavior. Boston: Graduate School of Business, Harvard University, 1967.
- Davis, H. Dimensions of marital roles in consumer decision-making. Journal of Marketing Research, 1970, 7, 168-177.
- Ferber, R., & Lee, L.C. Husband-wife influence on family financial behavior. Journal of Consumer Research, 1974, 1, 43-50.
- Gibbs, M. Decision-making procedures by young consumers. Journal of Home Economics, 1963, 55, 359-360.
- Granbois, D. H. Decision processes for major durable goods. In G. Fisk (Ed.), New essays on marketing theory. Boston: Allyn & Bacon, 1971. Pp. 172-205.
- Hempel, D. Family buying decisions: A cross-cultural perspective. Journal of Marketing Research, 1974, 11, 295-302.
- Hill, R., et. al. Family development in three generations. Cambridge, Mass.: Schenkman Publishing Company, 1970.
- Jaffe, L. J., & Senft, H. The roles of husbands and wives in purchasing decisions. In L. Adler and I. Crespi (Eds.), Attitude research at sea. Chicago: American Marketing Association, 1966. Pp. 95-110.
- Katz, E., and Lazarsfeld, P. F. Personal influence. Glencoe, Ill.: The Free Press, 1955.
- Sampson, J. M. Determinants of the employment status of the wife-mother. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, 1972.
- Wolgast, E. G. Do husbands or wives make the purchasing decisions? Journal of Marketing, 1958, 23, 151-158.



BEHAVIORAL DECISION MAKING:  
A COMPARISON OF THREE MODELS

J. Paul Peter  
Indiana State University

Lawrence X. Tarpey, Sr.<sup>1</sup>  
University of Kentucky

This study compared three alternative decision-making strategies evidenced in the consumer behavior literature. The first strategy, minimization of expected negative utility (perceived risk) was formulated as the probability times importance of six facets of potential loss; the second strategy, maximization of expected positive utility (perceived return) was formulated as the probability times importance of six facets of potential gain; the third strategy, maximization of expected net utility (net perceived return) was formulated as the difference between perceived return and perceived risk. Using brand preference scores for six brands of automobiles as the dependent variable, multiple regression analysis indicated that the net perceived return model could explain more variance and produce more significantly related variables than the other two models.

#### Introduction

The literature dealing with the topic of consumer behavior is replete with theoretical constructions and models of varying degrees of complexity. A careful review of this literature leads one to the inexorable conclusion that the search for a "grand theory" of consumer behavior would be quixotic. A more reasonable task for the scholar-researcher is to rigorously test some of the existing theoretical formulations as a means of discovering which ones are more perspicacious explanations of how specific purchase decisions are made. In this paper the authors attempt to do just this.

One class of consumer decision-making models can be appropriately labeled as "cognitive-rational" because they focus on the key organizing variables of risk and utility (or pay-off). As decision making strategies these models postulate a consumer who operates in a manner similar to the rational man of economic theory in that behavior is (a) goal directed, (b) calculated and (c) predicated upon some knowledge of the costs and benefits of alternative choices. From an operational standpoint these models can be described as additive utility models and the literature reveals three basic formulations (or strategies).

First, as evidenced by the work begun at Harvard by Bauer in 1960 and carried on by others, there is the "perceived risk" strategy which generally assumes that consumers act to minimize (or at least reduce) any expected negative utility associated with purchase behavior; no serious consideration is given to expected positive utility. Second, there are the so-called attitude models which focus on the benefits of products which are positively evaluated and there is little consideration of expected negative utility. Third, there are studies which emphasize the "valence" concept. These were pioneered by Lewin (1943) and Bilkey (1953, 1955) where the research recognized the fact that consumers perceive products as having both desirable (positive valence) and

undesirable features (negative valence). The implicit strategy in this research is that individuals attempt to maximize the "net valence" which is the arithmetic difference between expected positive and negative utility (i.e., "net perceived return").

Thus, in the context of a risk-return typology, there would appear to be three distinct strategies in terms of how consumers make decisions. These three strategies could be stated as:

- (1) Select the brand that minimizes expected loss-  
(perceived risk)
- (2) Select the brand that maximizes expected gain-  
(perceived return)
- (3) Select the brand that maximizes net expected gain-  
(net perceived return)

All three of these strategies have received some empirical support and two critical observations are relevant here: (1) to date no consumer behavior study has compared the three strategies in terms of their relative explanatory power and (2) few studies have investigated the various facets or dimensions of utility consumers consider in their selection process. With all this in mind the authors designed a study which formulated (based on previous literature) three models representative of each of the three strategies as a means of comparing them in terms of their relative explanatory power. Another purpose of the study was to investigate the facets of utility discussed in the consumer behavior literature which consumers consider in making certain types of purchase decisions.

In terms of the dimensions of perceived risk-return, six facets have been selected for use in this study. They are: (1) financial risk-return, (2) performance risk-return, (3) psychological risk-return, (4) physical risk-return, (5) social risk-return, and (6) time risk-return. Those particular facets have been selected because they were the most prominent and widely discussed in the perceived risk literature. It should be noted, however, that only one study (Jacoby and Kaplan, 1972) has discussed all six of these facets and argued that they are conceptually independent. Most studies have only dealt with a subset or, in some cases, combinations of these six facets.

### The Three Models

#### The Perceived Risk Model

Since the introduction of perceived risk to consumer behavior by Bauer (1960), the construct has been conceptualized as a dual-component, multi-faceted phenomenon. As defined by Kogan and Wallach (1964), the two components of perceived risk are "... 'a chance' aspect where the focus is on probability (of losing) and a 'danger' aspect where the emphasis is on severity of negative consequences." In terms of these components, Sieber, et al. (1964), Cunningham (1967), and Hansen (1972) have suggested a multiplicative model which could be described as:

$$PR = f (PL \cdot IL) \quad [1]$$

where PR = perceived risk  
 PL = probability of a loss  
 IL = importance of a loss

The above model is not a multi-faceted model and needs to be modified in order to accommodate the notion that consumer purchase risk is a multi-faceted concept. Thus, a multi-faceted model of perceived risk can be depicted as follows:

$$OPR = f \sum_{i=1}^n (R_i) \quad [2]$$

where OPR = overall perceived risk

$R_i$  = risk facets such as financial, performance, physical, psychological, social and time losses

Although discussed as a dual-component, multi-faceted phenomenon, the empirical research of perceived risk has been confined to measurements either of the facets or the components but seldom the two together. For example, the research done by Bauer, Cox, Cunningham, and others as Harvard, Cox (1967) as well as Copley and Callom (1971) and Hirsch, et al. (1972) employed general measures of both uncertainty and importance components but did not delineate the various facets of risk. On the other hand, Perry and Hamm (1969), Schiffman (1972), Roselius (1971), Jacoby and Kaplan (1972) and Zikmund and Scott (1973) ignored the components of perceived risk and used a general measure for each facet. Further, although Jacoby and Kaplan (1972) do not measure the two components, they suggested the following model for overall perceived risk:<sup>2</sup>

$$OPR = f (UFR \cdot CFR); (UPR_1 \cdot CPR_1); (UPR_2 \cdot CPR_2); \quad [3] \\ (UPR_3 \cdot CPR_3); (USR \cdot CSR) + \text{error}$$

where OPR = overall perceived risk

UFR = uncertainty of financial risk

CFR = consequences of financial risk

UPR<sub>1</sub> = uncertainty of performance risk

CPR<sub>1</sub> = consequences of performance risk

UPR<sub>2</sub> = uncertainty of physical risk

CPR<sub>2</sub> = consequences of physical risk

UPR<sub>3</sub> = uncertainty of psychological risk

CPR<sub>3</sub> = consequences of psychological risk

USR = uncertainty of social risk

CSR = consequences of social risk

Although they fail to hypothesize the relationship among the facets, in view of (1) the analogy drawn between perceived risk and the additive Fishbein Attitudinal Model (Zikmund, 1973), and (2) the consumer behavior additive utility models discussed by Sheth (1970) and Moinpour and McLachlin (1971), an additive model can be formulated to accommodate the need to deal with multi-risk facets:

$$OPR_j = f \sum_{i=1}^n (PL_{ij} \cdot IL_{ij}) \quad [4]$$

where OPR<sub>j</sub> = overall perceived risk for brand j

PL<sub>ij</sub> = probability of loss i from the purchase of brand j

IL<sub>ij</sub> = importance of loss i from purchase of brand j

n = risk facets

In this extended model the concept of perceived risk is depicted not only as a multiplicative function of probability of loss and importance of loss as in equation (1), but also an additive model of the various facets of risk as in equation (2).

### The Perceived Return Model

As pointed out by Wilkie and Pessemier (1973, p. 429), a basic formulation of the consumer behavior attitude model can be described as:

$$A_{jk} = \sum_{i=1}^n I_{ik} B_{ijk} \quad [5]$$

where  $i$  = attribute or product characteristic,  
 $j$  = brand,  
 $k$  = consumer or respondent

such that

$A_{jk}$  = consumer  $k$ 's attitude score for brand  $j$ ,  
 $I_{ik}$  = the importance weight given attribute  $i$  by consumer  $k$ , and  
 $B_{ijk}$  = consumer  $k$ 's belief as to the extent to which attribute  $i$  is offered by brand  $j$ .

Fishbein (1971, p. 315) refers to this model as an "attitude toward an object" model; however, the perceived return model used in this study is couched more in terms of an "attitude toward behavior" model. The importance of this distinction is given by Fishbein (1971, p. 315):

Thus, the salient beliefs I want to consider are beliefs about (the consequences of) buying or using the product rather than beliefs about (the attributes of) the product per se. . . . I think this distinction between attitude toward behavior and attitude toward an object is a crucial one, and one that has been overlooked. . . . All I am saying is that whether I buy product X . . . will depend more on my beliefs about (the consequences of) buying the product . . . than on my beliefs about (the attributes of) product X.

In addition to this argument, there are two other reasons why the "attitude toward an object" model was not used in this study. First, of prime importance for the "attitude toward an object" model is the specification of salient attributes; as Wilkie and Pessemier (1973, p. 432) have observed, "attribute specification is the weakest part of composition models." Also, although Wilkie and Pessemier (1973, p. 433) stated that in terms of attribute specification, "theoretical development is preferable," there is ". . . a lack of theoretical concepts for attributes." This problem coupled with the theoretical background in the perceived risk literature dealing with expected behavior outcomes, led logically to a preference for a perceived return model based on "attitude toward behavior."

Second, since a primary purpose of this study was to compare decision making models, it was felt that these models should be as compatible as possible. Consequently, comparing negative expected outcomes with positive product attributes was felt to be conceptually undesirable. Thus, the perceived return model was based on positive expected outcomes rather than on product benefits and can be depicted as:

$$OPR_{ej} = f \sum_{i=1}^n (PG_{ij} \cdot IG_{ij}) \quad [6]$$

where  $OPR_{ej}$  = overall perceived return for brand j  
 $PG_{ij}$  = probability of gain i from purchase of brand j  
 $IG_{ij}$  = importance of gain i from purchase of brand j  
 $n$  = return facets

Although the perceived return model was formulated identically to the perceived risk model except for the focus on positive utility (or gains) instead of negative expected utility, the two models are conceptually independent. In other words, the level of perceived risk has no necessary bearing on the level of perceived return.

### The Net Perceived Return Model

In essence, Lewin's vector hypothesis of consumer behavior states that (1) the net valence is the arithmetic difference between positive valences and negative valences and (2) if this remainder is positive, the purchase will tend to be made and vice versa. Based on this logic, the net perceived return model is formulated as:

$$NPR_{ej} = f (OPR_{ej} - OPR_j) \quad [7]$$

$$= f \sum_{i=1}^n [(PG_{ij} \cdot IG_{ij}) - (PL_{ij} \cdot IL_{ij})] \quad [8]$$

where  $NPR_{ej}$  = net perceived return for brand j  
 $PG_{ij}$  = probability of gain i from purchase of brand j  
 $IG_{ij}$  = importance of gain i from purchase of brand j  
 $PL_{ij}$  = probability of loss i from purchase of brand j  
 $IL_{ij}$  = importance of loss i from purchase of brand j  
 $n$  = utility facets

This model is a combination of the perceived risk and perceived return models. Its purpose is to test the maximization of net utility hypothesis. Conceptually, since this model takes into account (explicitly) both positive and negative expectations, it is intuitively the superior model.

### Method

Two questionnaires were employed in this study. Both questionnaires are identical except that questionnaire #1 contains items which deal with three brands of compact cars (Ford Pinto, Chevrolet Vega, Mazda RX3) while questionnaire #2 contains items dealing with three brands of intermediate cars (AMC Matador, Chevrolet Malibu, Volkswagen Dasher). Several different types of questions were asked. First there were questions relating to brand preference (the dependent variable). In addition, there were twenty-four items per brand intended to measure the independent variables: (a) six items intended to measure the probability of the six types of loss from purchase of the brand; (b) six items intended to measure the importance of the six types of loss from purchase of the brand; (c) six items intended to measure the probability of the



six types of gain from the purchase of the brand; and (d) six items intended to measure the importance of the six types of gain from the purchase of the brand. All items were scored on a seven point semantic differential scale (anchored "probable-improbable" and "important-unimportant") in order to make the data more interval-like (Moinpour and Wiley, 1972). Relative to the "gain" and "loss" facets an important point should be noted. Since no previous study had attempted to measure (1) both expected positive and negative utility and (2) both the probability and importance components separately for each of the six facets, the authors had to formulate operational definitions for this particular study. However, the definitions generated were based on conceptual discussions in the consumer behavior literature and operational definitions used in similar types of studies, primarily those of Fishbein (1967, 1971), Roselius (1971) and Jacoby and Kaplan (1972). All of these conceptualizations have face validity.

### Data Collection

Data were collected from a convenience sample of 217 juniors and seniors enrolled in Business Administration curriculum at the University of Kentucky. A total of 210 usable questionnaires were obtained. Of this total, 108 dealt with the compact brands and 102 dealt with the intermediate brands. Since this was an exploratory study to test the efficacy of three related models the authors felt that the problem of external validity was not critical; nevertheless replication of this study should use a population which would permit the results to be generalized.

The perceived risk, perceived return and net perceived return indices for each of the six brands were formed in the following manner. The perceived risk indices were formed by multiplying the appropriate responses to the probability and importance of loss questions; the perceived return indices were formed by multiplying the appropriate probability and importance of gain questions; and net perceived return indices were formed by subtracting each of the perceived risk indices from each of the perceived return indices.

### The Analysis

Eighteen (3 conceptual models x 6 automobile brands) multiple regression runs<sup>3</sup> were performed using the brand preference scores as dependent variables. Each computer run involved two procedures. First, each set of six indices were forced into the regression equation and regressed against the appropriate brand preference score. This procedure was employed to determine the total amount of variance explained ( $r^2$ ) by the six facets. Second, a stepwise procedure was employed which brought into the regression equation only those independent variables (facets) significantly ( $p < .05$ ) related to brand preference. This procedure was necessary to determine those facets of perceived risk, perceived return, and net perceived return which are significantly related to brand preference for each of the brands. In addition, examination of the beta coefficients for significant facets provides further insight into the relationships. Beta coefficients estimate the amount of change in brand preference which can be explained when a unit change in a facet is made while the T-ratio indicates whether this amount of change is significantly different than zero. Thus, the beta coefficients show the relative importance of each facet for each brand and the sign of the beta coefficient indicates whether the facet is directly or inversely related to brand preference. Conceptually, perceived risk should be inversely related to brand preference and perceived return and net perceived return should be directly related to brand preference; examination of the signs of the beta coefficients is a test of this notion.

## Results

Comparison of the Three Conceptual Models

Table 1 below shows all six brands studied along with the coefficients of determination for each of the three models. These  $r^2$ 's were generated by forcing all facets into the regression equation.

TABLE 1

Brand Preference as a Function of Perceived Risk, Perceived Return, and Net Perceived Return: A Comparison of  $r^2$ 's Forcing All Facets into the Regression Equation

Brand	Perceived Risk Model	Perceived Return Model	Net Perceived Return Model
<u>Compacts</u>			
Mazda	.2551 (.2104)*	.1034 (.0501)	.2237 (.1776)
Pinto	.3418 (.3027)	.1827 (.1341)	.3568 (.3186)
Vega	.0963 (.0426)	.2298 (.1841)	.2000 (.1525)
<u>Intermediates</u>			
Dasher	.2998 (.2556)	.2062 (.1560)	.3516 (.3106)
Malibu	.1471 (.0932)	.2863 (.2412)	.2937 (.2491)
Matador	.1034 (.0462)	.1393 (.0850)	.2066 (.1565)
Mean $r^2$ for Six Brands	.2072 (.1585)	.1912 (.1194)	.2720 (.2275)

\* ( ) indicates  $r^2$  adjusted for degrees of freedom.

In analyzing these data three points are noteworthy. First, for four of the six automobile brands (Pinto, Dasher, Malibu, and Matador) the net perceived return model explained more variance in brand preference than either the perceived risk or perceived return models. Second, for the two brands which the net model did not explain the most variance, the other models accounted for only one brand each and the net perceived return model in both cases explained the second largest amount of variance. Third, the average unadjusted and adjusted  $r^2$ 's were appreciably higher for the net perceived return model than for the other two models. Thus, if the amount of variance explained were the only available criterion for determining the validity of a model, based on this portion of the analysis, the net perceived return model appeared to be the preferred formulation.

Analysis of Significantly Related Variables--  
Perceived Risk Model

Table 2 below presents the perceived risk facets which were found to be significantly related to brand preference in the stepwise regression equation.

TABLE 2

Beta Coefficients, T-Ratios, and Significance Levels for  
 Perceived Risk Facets Significantly Related to Brand  
 Preference for Six Automobile Brands\*

Brand	Risk Facet	Beta Coefficient	T-Ratio	Significance Level
<u>Compacts</u>				
Mazda	Psychological	-.2347	2.6887	<.01
	Time	-.3771	4.3195	<.001
Pinto	Financial	-.2386	2.5051	<.025
	Performance	-.3944	4.1400	<.001
Vega	Financial	-.2701	2.8889	<.01
<u>Intermediates</u>				
Dasher	Financial	-.1866	2.1453	<.05
	Physical	-.4543	5.2237	<.001
Malibu	Time	-.3625	3.8895	<.001
Matador	Physical	-.2590	2.6816	<.01

\*Results reported here are final stepwise solutions.

Relative to the data in the above table three points should be noted. First, it is important to recognize that every brand had at least one risk facet which was significantly related to it. Second, all beta coefficients for the significantly related perceived risk facets were negative as they logically should be. Third, no single facet of risk dominated all six brands; this indicates that perceived risk may not be only product specific as Cunningham (1967) has shown, but also brand specific which is reasonable and consistent with the literature.

Analysis of Significantly Related Variables--  
Perceived Return Model

Table 3 is similar to the previous table. It presents the perceived return facets which were found to be significantly related to brand preference in the stepwise regression equation.

Relative to the data presented in Table 3 four points should be noted. First, it should be noted that every brand had at least one facet of return significantly related to it. Second, all beta coefficients for the significantly related perceived return facets were positive as one would logically expect.

TABLE 3

Beta Coefficients, T-Ratios, and Significance Levels for  
Perceived Return Facets Significantly Related to Brand  
Preference for Six Automobile Brands\*

Brand	Return Facet	Coefficient	T-Ratio	Significance Level
<u>Compact</u>				
Mazda	Performance	.2210	2.3337	<.025
Pinto	Time	.4153	4.7011	<.001
Vega	Performance	.3001	3.0136	<.005
	Psychological	.2125	2.1344	<.05
<u>Intermediate</u>				
Dasher	Psychological	.1880	1.9821	<.05
	Time	.3002	3.1647	<.005
Malibu	Financial	.3144	3.5261	<.001
	Psychological	.3364	3.7730	<.001
Matador	Psychological	.1921	2.0138	<.05
	Time	.2595	2.7192	<.01

\*Results reported here are final stepwise solutions.

Third, the fact that no single facet of return was significant for every brand suggests that perceived return may well be brand specific also which is not inconsistent with the theory. Fourth, none of the perceived return facets found significant were the same as the perceived risk facets found significant in Table 2. This suggests that consumers probably associate different dimensions of positive and negative utility with their brand preferences.

#### Analysis of Significantly Related Variables-- Net Perceived Return Model

Table 4 below is similar to the two previous tables. It presents the net perceived return facets which were found to be significantly related to brand preference in the stepwise regression equation.

Two points should be noted. First, the sign of all beta coefficients for the significantly related net perceived return facets were positive as one would expect if net perceived return and brand preference are truly directly related. Second, more net perceived return facets were found to be significantly related to brand preference than for either of the other two models. Twelve as opposed to ten for perceived return, and nine for perceived risk. Thus, if the validity criterion of the three models were based on the number of risk-return facets which were shown to be significantly related to brand preference, the net perceived return model would again be preferred.

TABLE 4

Beta Coefficients, T-Ratios, and Significance Levels for  
Net Perceived Return Facets Significantly Related to Brand  
Preference for Six Automobile Brands\*

Brand	Facet	Beta Coefficient	T-Ratio	Significance Level
<u>Compact</u>				
Mazda	Performance	.4272	4.8649	<.001
Pinto	Performance	.3843	4.2287	<.001
	Psychological	.2719	2.9917	<.01
Vega	Performance	.3903	4.3647	<.001
<u>Intermediate</u>				
Dasher	Financial	.2453	2.9023	<.01
	Psychological	.2424	2.6122	<.025
	Physical	.3114	3.3550	<.005
Malibu	Social	.2185	2.3947	<.025
	Psychological	.1900	1.9949	<.05
	Time	.3003	3.2359	<.005
Matador	Performance	.2498	2.4812	<.025
	Psychological	.2397	2.3814	<.025

\*Results reported here are final stepwise solutions.

#### Conclusions and Implications

The basic purpose of this study was to investigate three alternative decision making strategies in terms of their relative ability to explain brand preference. The three strategies, minimization of perceived risk, maximization of perceived return, and maximization of net perceived return were conceptualized, measured, and analyzed as multiple objective models and the results of these analyses were compared. Our findings have clearly indicated that the net perceived return model could explain more of the variance in automobile brand preference than the other two models and that perceived risk was a more potent explainer than perceived return. In addition, three other conclusions can be drawn from the results of this study. First, consumers consider expectations of both positive and negative utility in their automobile brand preference decisions. Second, perceived risk, perceived return, and net perceived return are brand specific rather than product specific. Third, automobile brand preference varies inversely with perceived risk and directly with perceived return and net perceived return.

Although there was both conceptual and empirical support for the variables and models employed in this study, more research is needed. The authors offer three suggestions for further verification and improvement of the net perceived return model. First, as Edwards (1961) has previously noted in a discussion of decision making models, the major problem is that the rules of combination

for the variables are not known, i.e., it is not known how consumers weight and combine variables to make decisions. Thus, although it was assumed in this study that probabilities and importances were combined in a multiplicative fashion, and that risk was subtracted from return to determine the net perceived return model, future conceptualization and research should investigate other methods of variable combination.

Second, although six facets of utility were determined for use in this study, there may be other important facets consumers consider. For example, expectations of gain and loss because of product obsolescence could conceivably be an important consideration. In addition, the social facet could perhaps be further disaggregated into expectations concerning friends, relatives, or other relevant reference groups.

Third, although this study attempted to replicate results using different samples, and different classes and brands of automobiles, much fruitful research could be done in further replication. In particular, different brands of automobiles should be studied. Different brands of other product categories should be considered. Wherever possible more reliable samples should be employed. Finally, replications should be done for different products and brands in different stages of the product life cycle although the present study attempted to do this by including two relatively new brands in the product sample, viz., the Mazda and the Dasher.

#### FOOTNOTES

<sup>1</sup>J. Paul Peter is an Assistant Professor of Marketing, Indiana State University, Terre Haute, Indiana, and Lawrence X. Tarpey is Professor of Business Administration, University of Kentucky, Lexington, Kentucky.

<sup>2</sup>They also mentioned "Since this study was conducted (Spring, 1970), Roselius (1971, p. 58) has identified a sixth variety of risk: Time loss. . ." (p. 393) and suggested that it should be included.

<sup>3</sup>K. J. Johnson, BIGSTEP Regression Program Version DVT4, University of Kentucky, Lexington, Kentucky, 1970.

#### REFERENCES

- Bauer, R. A. Consumer behavior as risk taking. In R. S. Hancock (Ed.), Dynamic marketing for a changing world. Chicago: American Marketing Association, 1960, 389-398.
- Bilkey, W. J. A psychological approach to consumer behavior analysis. Journal of Marketing, 1953, 17, 18-25.
- Bilkey, W. J. Psychic tensions and purchasing behavior. Journal of Social Psychology, 1955, 43, 247-257.
- Copley, T. P. and F. L. Callom. Industrial search behavior and perceived risk. In D. M. Gardner (Ed.), Proceedings. College Park, Maryland: Association for Consumer Research, 1971, pp. 208-231
- Cox, D. F. (Ed.). Risk taking and information handling in consumer behavior. Boston: Harvard University Press, 1967.

- Cunningham, S. M. The major dimensions of perceived risk. In D. F. Cox (Ed.), Risk taking and information handling in consumer behavior. Boston: Harvard University Press, 1967, 82-108.
- Edwards, W. Behavioral decision theory. Annual Review of Psychology, 1961, 12, 473-498.
- Fishbein, M. A. (Ed.) Readings in attitude theory and measurement. New York: Wiley, 1967.
- Fishbein, M. A. Some comments on the use of 'models' in advertising research. In Proceedings: Seminar on translating advanced advertising theories into research reality. Amsterdam, The Netherlands: European Society of Market Research, 1971.
- Hansen, F. Consumer choice behavior: A cognitive theory. New York: Free Press, 1972.
- Hirsch, R. D., R. S. Dornoff and J. B. Kernan. Perceived risk and store selection. Journal of Marketing Research, 1972, 9, 434-439.
- Jacoby, J. and L. B. Kaplan. The components of perceived risk. In M. Venkatesan (Ed.), Proceedings. Iowa City, Iowa: Association for Consumer Research, 1972, pp 382-393.
- Johnson, K. BIGSTEP Regression Program Version DVT4, University of Kentucky, 1970.
- Kogan, N. and M. A. Wallach. Risk taking: A study in cognition and personality. New York: Holt-Rinehart, 1964.
- Lewin, K. Forces behind food habits and methods of change. Bulletin of the National Research Council, 1943, 108.
- Moinpour, R. and D. L. Maclachlan. The relations among attribute and importance components of rosenberg-fishbein type attitude models: An empirical investigation. In D. Gardner (Ed.), Proceedings. College Park, Maryland: Association for Consumer Research, 1971, pp. 365-375.
- Moinpour, R. and J. B. Wiley. An approach to the resolution of multicollinearity in the attribute structure of attitudes. In M. Venkatesan (Ed.), Proceedings. Iowa City, Iowa: Association for Consumer Research, 1972, pp. 341-348.
- Perry, M. and B. C. Hamm. Canonical analysis of relations between socio-economic risk and personal influence in purchase decisions. Journal of Marketing Research, 1969, 6, 351-354.
- Roselius, T. Consumer rankings of risk reduction methods. Journal of Marketing, 1971, 35 (1), 56-61.
- Schiffman, L. G. Perceived risk in new product trial by elderly consumers. Journal of Marketing Research, 1972, 9, 106-108.
- Sheth, J. An investigation of relationships among beliefs, affect, behavioral intentions and behavior. Unpublished manuscript, University of Illinois, 1970.

Sieber, J. E. and J. T. Lanzetta. Conflict and conceptual structure as determinants of decision making behavior. Journal of Personality, 1964, 32, 622-641.

Wilkie, W. L. and E. A. Pessemier. Issues in marketing's use of multi-attribute attitude models. Journal of Marketing Research, 1973, 10, 428-441.

Zikmund, W. G. An empirical investigation of the multidimensional nature of perceived risk and related variables. Unpublished doctoral dissertation, University of Colorado, 1970.

Zikmund, W. G. and J. E. Scott. A multivariate analysis of perceived risk, self-confidence and information sources. In M. J. Schlinger (Ed.), Proceedings. Chicago: Association for Consumer Research, 1974.





THE USE OF SEALS OF APPROVAL IN CONSUMER DECISION-MAKING  
AS A FUNCTION OF COGNITIVE NEEDS AND STYLE

Thomas L. Parkinson<sup>1</sup>  
University of Delaware

The purpose of this study is to investigate the influence of two individual-difference variables, cognitive style and need for cognitive clarity, on the use of seals of approval in a consumer decision-making situation. Specifically, it is hypothesized that the use of seals as informational sources will be positively related to the individual's need for cognitive clarity, and that the Kelman-Cohler (1959) hypothesis concerning the interaction of cognitive needs and style will be in evidence.

To test these hypotheses a simulated shopping experiment was designed and administered to 198 adult female residents of Northern Delaware. During the course of this experiment each subject was asked to "purchase" four common household products from an array of unfamiliar but similar brands, some of which had a seal of approval and some of which did not. The choices made during the shopping experiment were then evaluated relative to the cognitive needs and styles of the subjects, and the hypotheses were supported.

The consumer decision process may be viewed primarily as being a problem solving activity involving the acquisition, processing, and transmission of information. During the decision-making process the consumer is exposed to informational cues from both personal and impersonal sources, each of which in turn may be classified as either advocative or independent depending on their predilection (Andreasen, 1965). Research finding concerning the role played by advocate personal sources (primarily salesmen), advocate impersonal sources (primarily advertisements), and personal independent sources (family, friends, and associates) in the consumer decision process, while not ample, may be found in the marketing literature. Included in this research are a number of studies which examine the influence of such individual-difference variables as cognitive needs and style on consumers' acquisition and use of information from these sources.

However, published research concerning the role of impersonal independent sources (neutral) of information has been largely limited to studies of the role of Consumer Reports (Beem, 1951; Beem and Ewing, 1954; Hempel, 1966; Sargent, 1959). Until recently, there has been virtually no published academic research, and only a limited amount of governmental and proprietary research concerning the role played by seals of approval in the consumer decision-making process (See Parkinson, 1972, pp. 4-24). Moreover, with the exception of a brief exploratory study by Cox (1967), there has been no research concerning the influence of various individual-difference variables on information acquisition from neutral sources.

It is the purpose of this study to investigate the influence of two individual-difference variables, cognitive style and need for cognitive clarity, on the use of seals of approval in a consumer decision-making situation.

### Cognitive Needs and Style

The results of Cox's (1967) study of the information handling of two consumers, and his follow-up study concerning information handling in making product evaluations (Cox, 1967a), suggest that the individual-difference variables, cognitive style and need for cognitive clarity are related to consumer information processing. Need for cognitive clarity can be considered a measure of the individual's need for certainty. Cognitive style refers to an individual's characteristic mode of resolving cognitive unclarity or uncertainty. Two types of cognitive style have been identified. One type, "clarifiers," have been found to react to unclarity by seeking new information and understanding. The other type, "simplifiers," attempt to reduce uncertainty by avoiding incongruous information thus simplifying their environment (Kelman and Cohler, 1959). Based on his initial findings, Cox (1967a) further investigated the effect of these variables on information handling, and found that consumers with a high need for cognitive clarity (certainty) were more likely to utilize available information when evaluating products. He was unsuccessful, however, in reconfirming the Kelman and Cohler result which attributes a reduced role to cognitive style when the need for certainty is low. However, Cox did observe this result in his exploratory research.

Based on these findings, the following hypotheses can be formulated:

1. In a consumer decision-making situation, the use of seals of approval as informational sources will be positively related to the individual's need for cognitive clarity.
2. In a consumer decision-making situation, the effect of cognitive style on the use of seals of approval as informational sources will be related to the individual's need for cognitive clarity as follows:
  - a. When need for cognitive clarity is high, cognitive style will have an effect on the use of seals with clarifiers making greater use of them than simplifiers.
  - b. When need for cognitive clarity is low, cognitive style will have no effect on the use of seals.

### Methodology

The two hypotheses were evaluated by means of a simulated shopping experiment. During the course of this experiment each subject was asked to "purchase" four common household products (an appliance, a dishwashing detergent, a child's undershirt, and a steak) from an array of unfamiliar but similar brands, some of which had a seal of approval and some of which did not.

The subjects taking part in the experiment were female adults drawn from a number of civic and church organizations in Northern Delaware which were selected with the objective of obtaining a representative cross-section of the population. The subjects were told, through their leaders, only that they would be participating in a marketing study. In return for their coopera-

tion, the groups were paid \$1.50 per participating member. In all, 198 subjects successfully completed the data gathering procedure.

### Measures

Use of seals and certifications as information sources. Since the consumer decision-making experiment involved the choice between similar products with and without various types of seals or certifications, the selection of a product bearing a seal or certification was interpreted as evidence that the individual had used the information provided by the symbol in making her decision. The validity of this interpretation was substantiated by means of a post-decision question concerning the reasons for the choices. This post-decision measure asked the subjects to indicate in their own words the reason for their selections during the experiment. In over 86% of the instances in which a seal-bearing product was selected, the presence of the symbol was given as the reason for the choice.

Cognitive style. Each subject was classified as either a "clarifier" or a "simplifier" based on her responses on the twelve item "Statements of Personal Preference Test" utilized by Cox (1967a). This measure consisted of a series of statements such as, "There is more than one right way to do anything" with which the subjects were asked to indicate whether they agreed or disagreed.

Need for certainty. Need for certainty was determined for each subject based on her responses to the eight item "Situational Response Test" previously used by Cox (1967a). This measure consisted of a series of statements concerning incidents involving cognitive uncertainty and the subjects were asked to indicate how frequently they would act in this manner. For example:

X is on a motor trip through the country. As evening approaches he finds himself in an unfamiliar area, lost and without maps or other guidance. He also finds that he is becoming terribly hungry. He decides to eat first and worry about finding his way later. You would act this way:

Always\_\_\_ Often\_\_\_ Sometimes\_\_\_ Once in a while\_\_\_ Never\_\_\_

### Procedure

The research was conducted in a number of church, civic, and university meeting halls in and around Newark, Delaware. The simulated shopping experiment was set up on a large table separated into four decision stations by a divider. At each decision station four brands of each product were displayed. In the cases of all products but the steaks, the brands included were comparable private label merchandise from outside of the Delaware Valley. Each brand was identified by a letter from A to P which appeared on small cards placed in front of each position. Three of the four brands at each decision station had one of the three seals shown in Table 1 while the fourth had no seal at all. Each seal thus appeared only once to each subject. Furthermore, the brand-seal combinations, and the positions of the products on the table were rotated throughout the experiment to minimize any interaction effect due to these factors.

TABLE 1

Seals Incorporated in the Choice Experiment by Product Category

Product Category	Seals		
Steak	U.S.D.A. CHOICE	Shop-N-Bag	Deluxe
Dishwashing Soap	Good Housekeeping	W. T. Grants	Approved
Undershirt	Parents' Magazine	Monsanto	Guaranteed
Hot Pot	Underwriters' Laboratory	Macy's	Tested

The subjects participated in the shopping experiment in groups of four. Following the verbal instructions of the experimenter, they were first asked to position themselves in front of one of the four decision stations, to inspect the products without touching them, to decide which one they would purchase, and then to mark their selection on the form provided. Then they were instructed to move clockwise around the table to the next decision station and to repeat the procedure for the next group of products. This process was repeated until the subjects had each made four "purchases."

Upon completion of the experiment, the subjects were asked to immediately sit down in an adjacent classroom area and to indicate on the bottom of their choice form the reasons for their selections. When this form had been completed, the subjects were asked to complete the cognitive style and need for certainty measures, and were excused.

#### Results and Discussion

In order to test the hypotheses, the subjects were first cross-classified with respect to their cognitive styles and need for cognitive clarity. A median split of the scores on the eight item "Situational Response Test" was made and each subject was classified as being either "high" (high scores) or "low" (low scores) in need for cognitive clarity. Then the subjects were classified with respect to their cognitive style based on their responses on the twelve item "Statements of Personal Preference Test." Again a median split of the scores was made with low scorers being classified as "simplifiers" and high scorers as "clarifiers."

Once the subjects were cross-classified in this manner, mean use scores for the subjects in each cell were calculated based on the number of seal-bearing products selected during the simulated shopping experiment and are shown in Table 2. Then using the same two-way classification, an analysis of variance with unequal cell sizes was carried out and the results are shown in Table 3.

TABLE 2  
Comparison of Use Scored as a Function of Cognitive Needs and Styles

Need for Cognitive Clarity	Cognitive Style		All Subjects
	Simplifier	Clarifier	
High	3.01 (45) <sup>a</sup>	3.97 (63)	3.57 (102)
Low	3.07 (35)	3.16 (55)	3.13 (90)
All Subjects	3.04 (80)	3.59 (118)	

<sup>a</sup>For example, there were 45 high need simplifiers, and their mean use score was 3.01.

TABLE 3  
Analysis of Variance Results:  
Use Scores as a Function of Cognitive Needs and Style

Source	d.f.	Mean Square	F-Ratio
Cognitive Style	1	12.935	13.74*
Need for Cognitive Clarity	1	11.711	12.44*
Interaction	1	9.048	9.61*
Error	194		

\*  $p < .05$

**Results.** Hypothesis 1 which examines the relationship between the subjects' use of seals of approval and their need for cognitive clarity is supported by the results of the analysis of variance. The variation in use scores resulting from differences in the subjects' need for cognitive clarity results in an F-ratio (12.44) which is significant at the .05 level. This result coupled with comparison of the mean use scores shown in the right hand column of Table 2, indicates that those subjects who indicated a high need for cognitive clarity made significantly greater use of seals and certification as informational sources than those low in this need, and Hypothesis 1 is therefore supported.

Hypothesis 2a and 2b can be evaluated by comparing the mean use scores in the center of Table 2. These results, coupled with the significant interaction indicated by the analysis of variance, give support to both of the hypotheses. As hypothesized, when the need for cognitive clarity among the subjects is high, cognitive style has a significant effect on the use of seals and certifications as informational sources with clarifiers using them

more. Therefore, Hypothesis 2a stands supported.

On the other hand, when the need for cognitive clarity is low, there is little difference in use scores relative to cognitive style and therefore, Hypothesis 2b concerning the role of cognitive style when need for cognitive clarity is low is also supported.

Discussion. These results basically confirm the Kelman-Cohler hypothesis concerning the interaction of cognitive needs and style in affecting the way people handle an ambiguous situation such as the choice experiment. It appears, even in the present experimental situation, cognitive clarifiers tend to utilize available informational cues and to act accordingly, whereas individuals who are classified as simplifiers tended to avoid this information. This effect seems to take place, however, only when the need for cognitive clarity is high: When this is not the case, the importance of dealing with the ambiguity is lower and cognitive style appears to play a reduced role.

#### Implications

The findings of this study that the use of seals of approval as informational sources in consumer decision-making is influenced by individual personality traits such as need for certainty and cognitive style would appear to have some limited implications for marketing strategy. Marketing managers should be aware that individuals with these personality characteristics will be more susceptible to influence by the presence of a seal or certification on a product than others. Therefore, when research exists concerning the presence of these personality characteristics in a particular target market that is of interest, marketing managers should give serious consideration to the incorporation of seals or certifications into the promotional mix for their product. However, identification of these individual traits in various market segments is difficult, and therefore, these findings may be difficult to operationalize.

#### FOOTNOTES

1. Thomas L. Parkinson is an Assistant Professor of Business Administration at the University of Delaware.

#### REFERENCES

- Andreasen, Alan R. Attitudes and consumer behavior: A decision model. In Lee E. Preston (Ed.), New research in marketing. Berkeley, California: Institute of Business and Economic Research, University of California, 1965.
- Beem, Eugene. Consumer-financed testing agencies. Journal of Marketing, 1951, 16, 272-285.
- Beem, Eugene, & Ewing, John. Business appraises consumer testing agencies. Harvard Business Review, 1954, 32, 113-126.

- Cox, Donald. The influence of cognitive needs and style on information handling in making product evaluations. In Donald Cox (Ed.), Risk taking and information handling in consumer behavior. Boston: Division of Research, Graduate School of Business Administration, Harvard University, 1967.
- Cox, Donald. Risk handling in consumer behavior--An intensive study of two cases. In Donald Cox (Ed.), Risk taking and information handling in consumer behavior. Boston: Division of Research, Graduate School of Business Administration, Harvard University, 1967. (a)
- Hempel, Donald. An experimental study of the effects of information on consumer product evaluation. (Doctoral Dissertation, University of Minnesota) Ann Arbor, Mich.: University Microfilms, 1966. No. 66-12208.
- Kelman, H. C., & Cohler, J. Reaction to persuasive communications as a function of cognitive needs and style. Paper presented at the meeting of the Eastern Psychological Association, Atlantic City, 1959.
- Parkinson, Thomas L. A study of seals and certifications of approval and their role in consumer information processing. (Doctoral Dissertation, University of Massachusetts) Ann Arbor, Mich.: University Microfilms, 1973. No. 73-6697.
- Sargent, Hugh. Consumer-product rating publications and buying behavior. University of Illinois Bulletin, 1959, 57, No. 31.





THE INFORMATION SEEKERS--A CROSS-CULTURAL CONSUMER ELITE<sup>1</sup>

Jack L. Engledow  
Indiana-Purdue University at Indianapolis

Hans B. Thorelli  
Indiana University

and Helmut Becker<sup>2</sup>  
University of Portland

This article summarizes an international comparative study of subscribers to and users of product test magazines in the United States and Germany. Such subscribers are found to be of higher income, education, and social class than "average consumers" in their respective countries, and are much more critical and "rational" consumers--with a strong interest in consumer affairs. They also, however, are somewhat more conservative and much more discriminating in their criticism of business and advertising than would be supposed by the usual consumerist stereotype. The striking similarity in demographic characteristics, attitudes, and purchase behavior across nations leads the authors to suggest that there exists a rather homogeneous cross-cultural elite of affluent and information-sensitive consumers which may be of strategic importance to consumerists, public policy makers, and marketing decision-makers.

The proper functioning of markets in advanced, affluent economies depends heavily upon adequate stocks and flows of information. As buying decisions become more complex, the quantity, quality, and type of information available to consumers substantially determines the allocation of resources, and becomes a critical concern for all who are interested in the efficiency and effectiveness of the marketplace.

One type of information which has grown rapidly in quantity, and presumably in influence in affluent economies, is the output of Consumer Information (CI) agencies--organizations whose purpose is to dispense information about products in which they have no commercial interest. These organizations, typified by Consumers Union in the United States and Consumers' Association in the United Kingdom, have grown precipitously during the past decade both in number and size of organizations; the number of consumers reached directly and indirectly by the information outputs of the agencies has increased dramatically.<sup>3</sup> The material which is summarized below is one facet of the International Consumer Information Survey--a five year study whose purpose was to gather an organized body of knowledge about CI agencies and the consumers which make use of their product information. Specifically, this portion of the investigation was designed to characterize and contrast the subscribers to three product test reporting magazines: DM and test in Germany, and Consumer Reports in the U.S.

Prior to this study, no comprehensive attempt had been made to identify users of product test reports and to study the ways and situations in which they used such information. There was ample evidence, however, that users--wherever found--tend to be much more affluent and better educated than the general population.<sup>4</sup> The consistency of this finding coupled with the over-

all purpose of the survey determined the following objectives of the subscriber study:

- 1). To characterize the subscribers of three major CI publications as to demographics and certain attitudes, and to gather data on their use of CI and other information sources.
- 2). To compare and contrast the subscribers of the three organizations, so as to isolate similarities and differences in subscriber characteristics among the different organizations.
- 3). To test the broad hypothesis that there exists a new breed of affluent, information-intensive consumers with consistent demographic and attitudinal patterns across cultural boundaries--a group to which we refer as "The Information Seekers."

### Methodology

The research described above can be characterized as a cross-cultural, comparative study of the subscribers to product test magazines. To fully meet the purposes of the study, it was necessary to make comparisons between subscribers and "average consumers" within countries and between groups of subscribers across countries. To accommodate these purposes and stay within reasonable budget constraints, the study was restricted to two countries, Germany and the United States and the multiple-sample research design described in Table 1. was used.

Six samples were selected as follows:

- 1) One nation-wide random sample of subscribers from each country for use in between-country comparisons. Over 600 usable mail questionnaires were received in each country from an initial mailing of 1000 (after two followups).
- 2) One metropolitan random sample of subscribers from each country for use in comparing with "average consumers." Professional market research firms administered a structured questionnaire to 100 subscribers.
- 3) One metropolitan random sample of "average consumers" from each country--drawn from the same area as the metropolitan subscribers, for within-country comparisons. Professional interviewers administered a structured questionnaire to 200 randomly-selected residents.

These six samples were administered a questionnaire which was basically the same for all samples. The mail questionnaires were shortened slightly by omitting a small section on convenience purchases, and two unaided recall questions were changed to aided. The German samples, of course, received a translated version; care was taken by means of pretesting, consultation with our German professional research firm, and careful supervision by the German national who is a member of our research team to assure compatibility in meaning between countries, and not just literal translation. There were also minor changes required because of differences in products, information sources, and demographic characteristic categories between countries. Overall, much care was taken to insure maximum comparability in form and meaning among the mail-interview versions

TABLE 1

## Summary Of All Samples International Consumer Information Survey

Name of Sample	Country	Sample Size	Nature of Sample	Purpose of Sample
Nationwide Mail Subscribers Sample	U.S.	630	Systematic random sample of subscribers to <u>Consumer Reports</u> .	Intercountry comparisons on subscriber characteristics, behavior, attitudes, and perceptions of agencies and information. Intra-country analysis of satisfaction in purchases using agencies' information vs. those not using, etc.
	Germany	610	Systematic random sample of subscribers to <u>DM</u> and <u>test</u> .	
Metropolitan Interview Random Sample	U.S.	200	Stratified random sample of Indianapolis residents.	Intracountry comparisons with interview subscribers. Analysis of non-subscriber use of CI. Analysis of information use and needs of "average" consumer.
	Germany	198	Stratified random sample of Frankfurt residents.	
Metropolitan Interview Subscribers Sample	U.S.	101	Systematic random sample of Indianapolis subscribers to <u>Consumer Reports</u> .	Intracountry comparisons with interview random sample. More comprehensive and accurate recording of behavior, attitudes, perceptions and demographic characteristics than by mail survey.
	Germany	97	Systematic random sample of Frankfurt subscribers to <u>DM</u> and <u>Test</u> .	

and the German-American versions. Two versions of each style questionnaire were prepared, with questions split and reversed where appropriate to avoid positive-negative bias by respondents.

The questionnaires contained seven sections: General Attitudes; Durable Good Purchase Behavior; Auto Purchase Behavior; Convenience Good Purchase Behavior; Media Usage and Advertising Attitudes; CI Usage and Consumer Concerns; and Demographic Characteristics. The questionnaire was designed with attitude and purchase behavior questions in the beginning, before any overt reference to CI agencies, and care was taken in questionnaire design and interviewer training to mask the purpose of the study. (See Table 2). Purchase behavior questions were based upon the respondent's most recent purchase in the product category, and those with no purchase during the past year skipped the section altogether. Preliminary versions of all questionnaires were pretested, and after revision, were administered during the spring of 1970.

Analysis consisted primarily of two-group comparisons between subscribers and average consumers within countries, and between subscriber groups between countries. Tests for statistical significance were performed, using the Mann-Whitney U Test where ordinal data were available, and the Chi-Square Two Sample Test when the data were nominal.<sup>5</sup>

### Results

As illustrated in Figure I, the analyses in this study unfolded sequentially, starting with subscriber--average consumer comparisons of metropolitan samples in each country, followed by comparison of nation-wide samples of subscribers between countries.

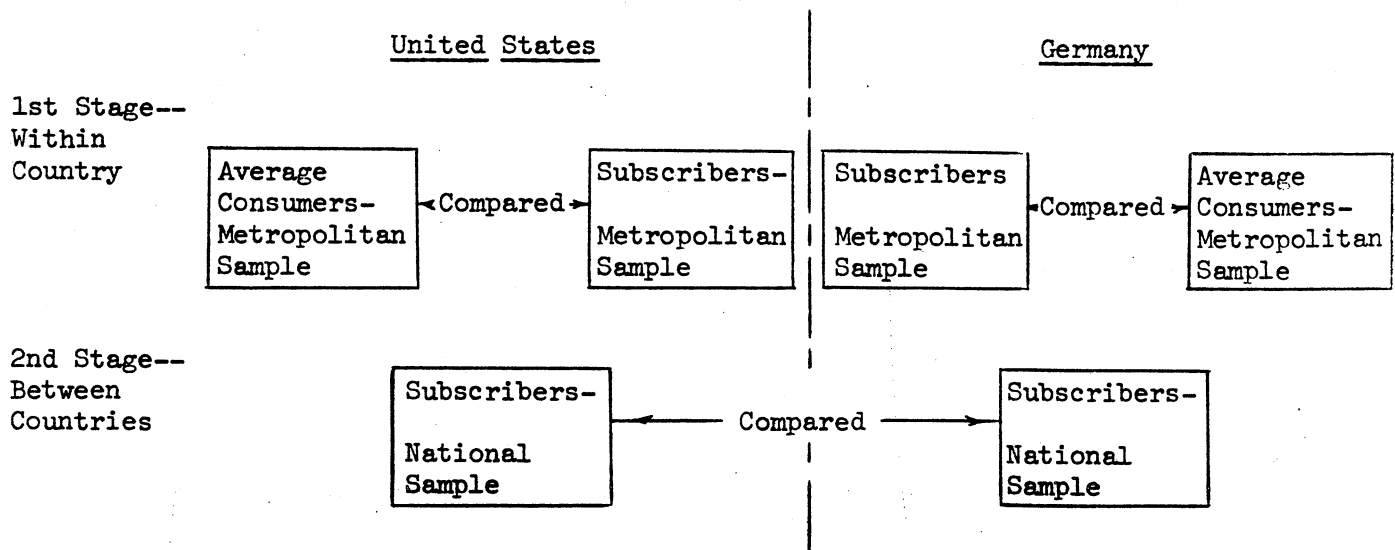


Figure 1. Description of comparisons in study

In brief, the first major question was how do subscribers compare to average consumers in their own country?; the second was how do subscribers in one country compare to those in another? These will be examined in turn:

TABLE 2

## Master Questionnaire Makeup by Sections

Questionnaire Section	Content
Section 1: General Atti- tudes	Seventeen Likert-scale questions on attitudes-- planfulness, liberal-conservative, opinion leader, business and government.
Section 2: Durables Purchase Data	Inventory of brands, types owned and date of pur- chase from a list of fifteen durable products. Specific purchase data on latest purchase, including planning, experience, specific concerns in shopping, information sources, role in purchase, shopping activity, and satisfaction with product, information sources, and own shopping effort, also forced brand preferences for five durable products.
Section 3: Auto Purchase Data	Inventory of autos owned plus purchase data on most recent auto exactly as in Section 2.
Section 4: Convenience Purchase Data	Inventory of brand preferences for all "frequently purchased" brands from a list of fifteen products plus experience data and the satisfaction sequence as in Sections 2 and 3.
Section 5: Media Usage Data	Information on use of radio and TV, magazine reader- ship, readership of ads in these media, and of direct questions on attitude toward advertising.
Section 6: Consumer Infor- mation Data	Rankings of priorities on series of consumerist issues and preferences on government activity in the area, comparison between advertising and comparative testing as an information source. Series of ques- tions on subscription, readership, usage, and satis- faction with product test magazines.
Section 7: Demographic Characteristics	Series of questions on demographic characteristics, including sex, age, income, education, occupation, political party, housing, leadership positions, marital status, and membership in union.

Within-Country Comparisons

Out of the hundreds of paired comparisons performed between the subscribers and average consumers in the two countries, a number of strong patterns emerge:

Demographic Characteristics. The previous finding that users of CI are affluent and well educated was strongly supported. In both countries, subscribers were found to have much higher incomes and educational levels, and were higher in social class than average consumers. They were also likely to be married, are slightly older in mean age (with heavy concentration in the 24-44 age bracket), and are much more likely to be in the professional and managerial occupational brackets. A summary of actual results making the same point with the national samples is included under Between Country Comparisons below.

General attitudes. There is also a striking pattern of similarity among subscribers in the two countries in a number of purchase-related attitudes which were measured by the survey.

TABLE 3

## General Attitude Comparisons

<u>Attitude Dimension</u>	<u>Highest Group</u>	
	<u>Germany</u>	<u>U.S.</u>
Planfulness	Sub*	Sub*
Opinion Leadership (self designated)	Sub*	Sub*
Early Adopter	AC	AC
Toward Business	Sub*	Sub*
Liberal-Conservative	Mixed	Mixed

Sub = Subscriber; AC = Average Consumer

\* Significant at .05 - Mann Whitney U Test.

Subscribers were found to be more planful, more likely to be opinion leaders, with a more favorable attitude toward business, and less likely to be product innovators than average consumers. The liberal-conservative dimension, however, was decidedly culture-linked as we measured it, with German respondents overall much more sympathetic with liberal social and economic actions than Americans. There were no systematic cross-cultural differences between subscribers and average consumers here, though on the general liberalism screening question "You can't change human nature," subscribers in both countries were more liberal.

Media usage. In both countries, subscribers watched TV and listened to radio less, and read newspapers and magazines more than average consumers. As would be expected from the education-income profile, subscribers showed higher readership of high brow magazines such as New Yorker and Capital (Germany), and less for the "lower brow" Modern Romances and Neue Revue. There were also wide differences in radio-television weeklies (TV Guide, Hoer Zu!) where average consumer readership was much higher, and in other product opinion magazines (Mechanics Illustrated, Auto-Motor Sport (Germany), Good Housekeeping) which were read more by subscribers to product test magazines. Besides the usual income-education bias in media selection, there seems to be strong correlation in readership of magazines containing product information. Another finding of interest was that the product test magazines were well known to about 20% of non-subscribers in each country, and that there is much evidence that product ratings and other information influence non-subscribers by means of borrowed copies, library usage, and word of mouth.

Purchase behavior. Subscribers own more autos and other durable goods than average consumers in both countries. Respondents indicated ownership from a list of fifteen of the most common durable goods items in each country. For all but one item (Black and white TV in Germany), subscribers were more likely to be owners, and they owned almost 1.5 times as many autos. This fact, combined with the opinion leadership finding above suggests that subscribers constitute a market segment whose economic importance may be disproportionate to its absolute size.

Respondents were asked to indicate perceived importance of various shopping criteria and information sources in the purchase of a specific durable good or auto. With thirteen shopping criteria, eight information sources, two types of purchases, and four samples in two countries, the data are obviously unwieldy to summarize, yet three fairly clear patterns were apparent:

1. Subscribers regard themselves as more "rational" consumers in the traditional sense: Compared to average consumers, they consistently perceive themselves as giving greater weight in buying to performance--economic related criteria such as durability, service, price, and performance; and less to shopping convenience variables like dealer location and product availability. Operational economy was important to German subscribers, but not to American subscribers in this pre-energy crunch survey. We could guess that the answer would be different today.
2. Subscribers use more sources of information in making purchases, and they rely heavily upon the product test magazines. In the U.S. over one-third of the subscribers cited Consumer Reports as an important information source (using unaided recall before mentioning Consumer Reports in the interview)--more than any other information source mentioned. Despite their experience as shoppers and their high educational level, subscribers perceive themselves as using past experience and personal observation less than average consumers.
3. Subscribers cited Information Availability as an important shopping criterion in 75% of the shopping experiences examined in both countries, as opposed to about 50% in the case of average consumers. This is but another indication of the importance of information to this special group.

Satisfaction. The subscribers examined here have spent money to receive specialized product information, and apparently a large percentage spent the time to sort through the detailed and often-complicated findings included in the magazines. (See the perceived usage data under Between Country comparisons below). One interesting question is whether such expense and effort results in "better" buying behavior from the user's viewpoint. At least one recent study would suggest that increased effort should lead to higher satisfaction.<sup>6</sup>

Satisfaction measures were taken from respondents on each purchase traced (durable, auto, and convenience) on three separate dimensions: satisfaction with the product purchased (e.g. "was all I expected"), with personal shopping activity ("could not have made better purchase, even with more information") and with information availability ("had plenty of information available").

A direct comparison of subscribers with average consumers in both countries revealed one interesting difference: German subscribers were generally better satisfied than average consumers with their overall buying behavior by the measures used, while U.S. subscribers were generally less satisfied. (See Table 4).



TABLE 4

Summary of Selected Satisfaction Measures  
(3 purchase decisions, 6 questions on each)

	<u>Higher Satisfaction Subscribers</u>	<u>Higher Satisfaction Average Consumer</u>
German	16	2
United States	5	13

If this meant that subscription to Consumer Reports led to poorer purchasing, it was a disappointing outcome (particularly to Consumer Reports). Further analysis clarified the situation. Within the U.S. subscriber group, it was possible to separate those who perceived themselves as having used CR as an important input in the particular purchases examined from those who did not. When these groups were compared, perceived users were higher in satisfaction for 15 of the 18 measures. The implication is that while subscribers as a group are harder to satisfy than average consumers, they do increase their satisfaction in those particular instances where they perceive that Consumer Reports was an important input into the decision. Overall, U.S. subscribers thus seem to either set higher standards for products, or to evaluate performance more critically, or both. The interesting difference between German and U.S. subscribers noted here will be discussed more thoroughly in Between-Country Comparisons below.

Summary. Out of the many findings in the within-country comparisons, it is particularly interesting given the purposes of this study to isolate variables in which both German and U.S. subscribers varied significantly from average consumers in their respective countries. For convenience, these may be divided into "Universals"--variables where German and American subscribers varied from average consumers in the same direction; and "Opposites"--where subscribers differed in each country, but in opposite directions:

Universals

Subscribers Higher than Average Consumers

Socio-economic

Income  
Education  
Social Class

Attitudes

Planfulness  
Opinion Leadership  
Opinion of Business

Buying Behavior

Ownership of Durables  
and Autos  
Purchase Experience  
Shopping Activity  
Concern with Information  
Availability

Media Usage

Print Media Readership  
Product Tests as Information Source

Other

Subscribers generally more favorable to increased government activity in the several action areas listed (product testing, preventing misleading advertising, etc.)

Universals (continued)

Subscribers Lower than Average Consumers:

Media Usage

Broadcast Media Usage (Radio-TV)

Opposites

German Subscribers Higher; U.S. Lower:

Attitude

Increased Government Control of Business

Satisfaction

All Satisfaction Measures.

U.S. Subscribers Higher; German Lower:

None

Selected ones of these Universals and Opposites will be examined in greater detail in the Between-Country Comparisons.

Between-Country Comparisons

The Within-Country comparisons suggested that subscribers in each country are high-income, well-educated, upper-middle social class consumers who are information sensitive, "rational," and--at least in the U.S.--hard to please. Comparisons of the nation-wide random sample of subscribers from the two countries added detail to this picture, and also gave a better opportunity to examine use of ratings and perceived effectiveness of the product testing organizations themselves.

Demographics. In Table 5, the basic demographic characteristics of subscribers are compared to census data. It is apparent not only that subscribers are sharply different from average consumers in each country, but also that they are quite similar between countries. Although no statistical tests were performed because of the differences in measures of income and education, it is apparent that subscribers are similarly over-represented in the highest income group, the highest educational group, the managerial-professional group, and the 25-44 age group. These findings have been consistently confirmed in other countries, as well.<sup>7</sup>

Attitudes. Attitudinal patterns for German and American subscribers were quite similar for most dimensions tested. Both groups were highly "planful," perceived themselves to be opinion leaders but not product innovators, and had moderately favorable attitudes toward business, as in the within-country comparisons. In the liberal-conservative area, however, the cultural differences between the two countries became obvious. The U.S. subscribers were substantially opposed to three liberal "action" areas--(increased government control of business more student power, welfare)--while the Germans strongly favored such policies.

Attitudes toward advertising in the two countries also present an interesting pattern: Americans are everywhere more favorable to advertising than Germans; but within-countries, American subscribers have a much less favorable attitude than average consumers, while German subscribers have a slightly more favorable attitude. Perhaps even more significant is the fact that while German criticism

TABLE 5

Selected Demographic Characteristics  
Subscribers Compared to Census Data

	<u>United States</u>		<u>Germany</u>	
	<u>Subscribers</u>	<u>Census</u>	<u>Subscribers</u>	<u>Census</u>
<u>Income</u>				
Over \$15,000	47%	22%	56%	15%
10,000-14,999	33	27	29	25
5,000-9,999	18	32	13	38
Under 5,000	2	19	2	21
	<u>100%</u>	<u>100%</u>	<u>100%</u>	<u>100%</u>
<u>Education</u>				
College Graduate (16 yrs.+)	40%	11%	46%	5%
Some College (13-15)	20	10	22	17
High School Grad. (12)	20	34	26	43
Some High School (9-11)	7	17		
Grade School of Less (8-)	3	28	6	32
	<u>100%</u>	<u>100%</u>	<u>100%</u>	<u>98%</u>
<u>Occupation</u>				
Professional and Technical Managers and Proprietors	37%	15%	35%	NA
Clerical	4	7	24	24%
Sales	6	18		
Craftsmen & Foremen	11	14	12	21
Operatives, Service & Laborers	10	36	10	26
Misc.	8	0	13	20
	<u>100%</u>	<u>100%</u>	<u>100%</u>	<u>100%</u>
<u>Age</u>				
Under 25	5%	18%	8%	12%
25-34	31	19	34	21
35-44	23	17	24	18
45-54	22	17	17	15
55-64	14	14	11	16
65 and Over	5	15	6	18
	<u>100%</u>	<u>100%</u>	<u>100%</u>	<u>100%</u>

is "across the board," American respondents--particularly subscribers--express strong support for the economic aspects of advertising (is essential, raises standard of living), but are highly critical of what might be called social aspects (persuades people to buy unneeded products, does not present a true picture). The pattern again suggests that subscribers are critical of some business practices, but basically sympathetic with the business system.

### Use of Product Test Reports

Table 6 summarizes the subscribers' perceived use of product test recommendations.

TABLE 6

Perceived Use of Product Test Recommendations  
National Mail Samples  
(As % of Actual Purchases)

	U.S.		German	
	<u>Durable</u>	<u>Auto</u>	<u>Durable</u>	<u>Auto</u>
Aware of Recommendations	64%	47%	56%	27%
Recommendations Important in Purchase Decision	53	38	52	18
Purchased Highly Rated Brand	41	29	45	18
Agree with recommendations after experience with products? Index <sup>a</sup>	92	86	89	89

<sup>a</sup>Index = (% of sample strongly agree or agree minus % sample disagree or strongly disagree)

Though such perceptions may be overstated, the results show three trends that were consistent wherever measured in the study:

- 1). Subscribers perceive high usage of recommendations in making purchases. This was shown by unaided recall in the metropolitan samples besides direct questioning in various parts of the mail and interview surveys.
- 2). Use is higher in durables than for autos in both countries. Coverage of autos is spotty in the German magazines, but extensive in Consumer Reports.
- 3). There is high satisfaction with the magazines' recommendations by those who have purchased recommended products.

The between-country comparisons also added another dimension to the findings on satisfaction which were cited in the Within-Country section above. There it was noted that German consumers were more satisfied with purchases overall than average consumers, U.S. consumers less satisfied, but that every case of perceived use of product test reports in an actual purchase was accompanied by higher

satisfaction with the product and the purchase process. These findings were strongly reinforced by the between-country comparisons. In the twelve possible comparisons (two products, three satisfaction dimensions, two questions each) German subscribers perceived themselves as better satisfied than Americans in all twelve. Either Germans are easier to satisfy or German information and products are superior. In comparing purchase decisions in both countries where there was perceived use of product tests, there was greater satisfaction in 20 of 25 measures--10 at statistically significant levels. None of the five measures where satisfaction was less was at a significant level.

Satisfaction with magazines and test reports. The subscribers to the product test magazines surveyed gave high performance marks to the magazines and the agencies which produce them. Figure 2 summarizes the results of twelve separate questions which were designed to measure three basic dimensions: Reliability-Credibility; Relevance-Timeliness; and Clarity. Two questions also measured general effectiveness.

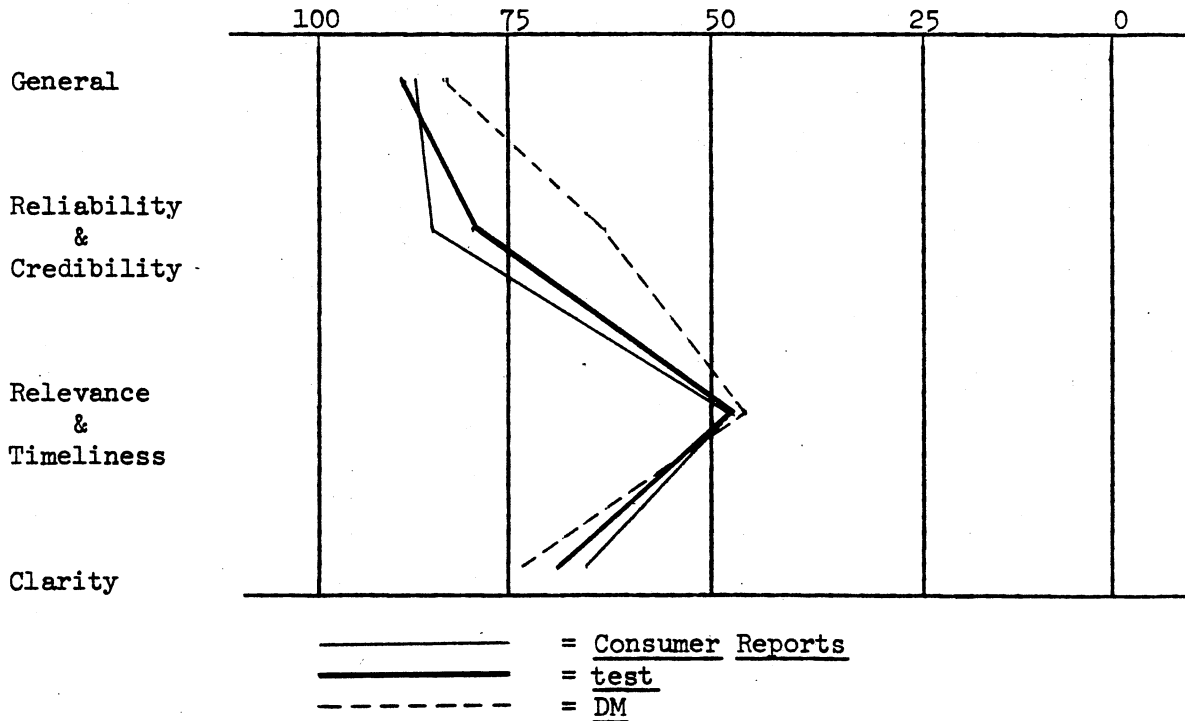


Figure 2. Satisfaction with magazines and agencies on four dimensions

The results are presented as a mean index, where 100 indicates all positive responses, minus 100 all negative. Though results are obviously highly favorable to the magazines, two interesting conclusions may be drawn:

- 1). DM, which is a commercial, comparatively flamboyant, popular style magazine which accepts advertising, gets higher ratings for clarity, but much lower marks on reliability and credibility than CR and test, both of which are rather straight-laced, non-profit, and assiduously avoid any commercial contact with business whatever.
- 2). All three magazines get the most criticism in Relevance and Timeliness--not testing often enough, testing too few brands or products, irrelevance of testing criteria (too much emphasis on safety in particular), and difficulty in finding recommendations at time of purchase.

It should be noted that these are relative weaknesses only, and that overall evaluation of magazines and agencies by subscribers was highly favorable throughout the study.

### Conclusions and Implications

This study has taken an extensive cross-cultural look at a comparatively small, rather specialized group of consumers--those who subscribe to product test magazines. Some cross-cultural differences were noted among subscribers. The Germans, with a long history of government paternalism, were more sympathetic to government action on most issues. They also are a bit behind the Americans in the overall game of consumption, so they own fewer goods, have less purchase experience, and seem not to be as hard to please (despite their stereotyped reputation for being methodical and rational). Overall, however, subscribers proved to be a surprisingly homogeneous group who might be characterized as follows:

They comprise a consumer elite of high income and educational attainment, tending to be upper middle class managers and professionals, concentrated in the 25-44 age bracket. They are hard-to-please consumers, critical of advertising and some business practices, and anxious for more government activity in various consumerist areas. At the same time, their criticism is selective, and in other areas they show conservative leanings and basic sympathy with business and private enterprise. They own many goods, are experienced and vigorous shoppers who plan carefully, and they use "rational" (economic-performance related) shopping criteria. They are highly information-intensive, seeking out and using many different types of product information,<sup>9</sup> and are particularly sensitive to the availability of "good" information in any given purchase situation. They use product tests extensively, believe in them and the agencies which dispense them, and are better satisfied when they use their recommendations. They perceive themselves to be opinion leaders, but not innovators in the use of products.

We have labeled this group "The Information Seekers." Some results in this study and others suggest that consumers with these characteristics may be found in all advanced, affluent economies, both within and without the ranks of subscribers to product test magazines. Common sense would dictate, and beginning evidence suggests, that some subscribers are not Information Seekers, and some Information Seekers are not subscribers, though the correlation between these two groups is probably high.

Subscribers to test magazines seldom comprise more than 3% of the households in a country (Norway is an exception) and there is insufficient data to suggest how much larger, if any, the Information Seeker group is likely to be. What is apparent is that from the standpoint of both sheer economic potential and capacity to influence others, this group probably has importance far beyond its actual numbers. It should delineate an information-sensitive market segment of some importance to marketing decision-makers in the private sector. Perhaps even more importantly, from the public viewpoint, the Information Seekers may well contain the vigilantes of the modern market place--well-informed, interested, and

vocal opinion leaders who serve as spokesmen and proxy purchasing agents for other consumers. In either role, they are an important group in our economy. They merit the further attention of both researchers and policy makers.

#### FOOTNOTES

1. This research was made possible by the generous support of the International Business Research Institute of the Graduate School of Business, Indiana University; and the Consumer Research Institute.
2. Jack L. Engledow is Associate Professor of Marketing at Indiana-Purdue University at Indianapolis; Hans B. Thorelli is the E. W. Kelley Professor of Business Administration at Indiana University; Helmut Becker is Assistant Professor of Marketing at the University of Portland.
3. The growth and nature of CI organizations is examined in detail in H. B. Thorelli and S. V. Thorelli, Consumer information handbook: Europe and North America. New York: Praeger Publishers, 1974.
4. This has been consistently reaffirmed by the annual subscriber surveys of Consumers Union, Consumers' Association, and other organizations in the U.S. and Europe, as well as in several independent surveys of such organizations.
5. See S. Siegel, Non-parametric statistics, New York: McGraw-Hill, 1956.
6. Richard N. Cordozo, An experimental study of consumer effort, expectation, and satisfaction. Journal of Marketing Research, Aug 1965, 2, 244-249.
7. Subscriber Surveys of Which, the publication of Consumers' Association in the United Kingdom, and a special survey in 1969 of subscribers to the Norwegian publication Forbruker-rapporten, for example.
8. The questions used in this section were identical to those used in a section of R. A. Bauer and S. A. Greyser, Advertising in America: The consumer view, Boston, Mass.: Division of Research, Graduate School of Business Administration, Harvard University, 1968. Results in the U.S. were similar to those obtained by Bauer and Greyser.
9. See H. B. Thorelli, Concentration of information power among consumers. Journal of Marketing Research, 8, 1971, VIII, 427-32.

#### REFERENCES

- Bauer, R. A., & Greyser, S. A. Advertising in America: The consumer view. Boston: Division of Research, Graduate School of Business Administration, Harvard University, 1968.
- Becker, H. Consumer information and the image of advertising in Germany with significant comparisons to America. Unpublished doctoral dissertation, Graduate School of Business, Indiana University, 1971.
- Engledow, J. The Consumer Reports subscriber: Portrait of an intense consumer. Indiana Business Review, Aug-Sept 1972, 32-40.

- Engledow, J. The impact of Consumer Reports ratings on product purchase and post-purchase product satisfaction. Unpublished doctoral dissertation, Graduate School of Business, Indiana University, 1971.
- Thorelli, H. B. Concentration of information power among consumers. Journal of Marketing Research, 8, 1971, 427-32.
- Thorelli, H. B. Consumer information programmes. International Consumer, Autumn 1972, 15-21.
- Thorelli, H. B. Testing, labeling, and certifying: A new perspective on consumer information. British Journal of Marketing, Autumn 1970, 126-32.
- Thorelli, H. B., & Thorelli, S. V. Consumer information handbook: Europe and North America. New York: Praeger Publishers, 1974.
- Thorelli, H. B., Becker, H. & Engledow, J. The information-seekers--an international study of consumer information and advertising image. Cambridge, Mass.: Ballinger, in press.





A GENERALIZED UTILITY MODEL OF SHOPPING BEHAVIOR<sup>1</sup>

Jimmy E. Hilliard<sup>2</sup>  
University of Georgia

Ronald L. Vaughn<sup>2</sup>  
Bradley University

and Fred D. Reynolds<sup>3</sup>  
University of Georgia

Since the 1920s there has been a continuing interest in building and testing formal models of consumer spatial behavior. This paper continues the model building tradition by developing a generalized model of spatial behavior and by presenting a preliminary test of the model in the context of intraurban clothing shopping.

Consumers engage in many decisions and activities in many different settings and situations. One almost inescapable activity, however, is spatial mobility -- an acknowledgement to Detroit. In the 1930s and 40s the auto liberated the shopper. And, free to roam, shoppers have traveled to the stores, shopping centers, and even cities that offered them what they wanted when they wanted it.

Spatial mobility, geographical population redistribution and other consumer and competitive factors have caused the framework of retailing to maintain an almost constant state to flux. To survive and prosper, retailers must attract shoppers and convert them to buyers. To do so requires, in part, an understanding of consumer spatial behavior or at least the ability to predict that behavior -- the subject of this paper.

#### The Situation

The opening of Northgate, a regional shopping center in Seattle, in 1950 launched a retailing innovation that rapidly diffused to other metropolitan suburbs during the 1950s and 60s. Given new shopping alternatives, consumers changed their spatial behavior and it became necessary for students of that behavior to explore new bases for predicting it. Some researchers tried Reilly's formula, substituting square footage of selling space and driving time for population and distance. Generally, however, the results were mixed and unsatisfactory. Other researchers have followed the lead of Huff who formulated a spatial model based on Luce's choice axiom (Huff, 1962; and Luce, 1959).

Huff's central thesis is that consumer spatial behavior is best described as a probabilistic phenomenon, which can be modeled in terms of consumers' perceived utilities or benefits of alternative shopping centers. That is, the probability of a consumer choosing a given shopping center is equal to the ratio of the utility of that center to the combined utilities of all centers under consideration.

### Some Problems

The spatial model developed by Huff has several appealing features: (1) it is based on a theory of individual choice behavior; (2) it incorporates variables that are readily obtainable, and (3) it can be applied in a straightforward manner to predict intraurban choice and to calculate sales potentials for existing and proposed center locations. Nevertheless, there are several problems related to the full utilization of the model including appropriate measures of utility, appropriate parameter estimation procedures, population heterogeneity, and inaccurate prediction in certain situations. Only the latter is examined in this paper.<sup>4</sup>

Huff noted in his pilot test that the model consistently obtained poorer predictions in situations where more than one shopping center were in close proximity. This problem may arise because the model assumes independence between centers. That is, the model does not allow for the perceived utility of one center to be affected by its position relative to another center -- a notion inconsistent with the concepts of complementary and substitute centers.

This paper attacks the problem of center independence by presenting preliminary findings of a comparative study of the Huff model and a more general utility model that allows for any existing interaction effects of centers to affect the calculation of consumers probabilities of shopping center choice. The test obviously is limited in that the other problems associated with operational choice models are ignored.

### The Generalized Utility Model

In thinking of generalized utility, consider first the somewhat analogous n-body problem in planetary attraction. Here the force vector in a given direction acting upon a point mass depends not only on the mass and distance of the body lying along that vector but upon the mass (and distance) of all surrounding bodies. Specifically, in the generalized n-body model, forces generated by all bodies are summed in vector fashion to yield the direction and magnitude of the net force vector and from this components of force in any direction may be obtained. Thus the concepts of planetary attraction are consistent with interdependence among surrounding bodies. We do not propose to represent the shopping process as being completely analogous to planetary attraction. However, consideration of the n-body model suggests an important generalization of the two-body model as posited by Huff.

In developing this model we denote first the intrinsic drawing power of center j by  $I_j$ , assumed to be independent of other centers. However, a center's net drawing power potentially is enhanced by substitute and complementary contributions from neighboring centers. Also, the assumed contribution is attenuated by the disutility or cost between centers. Thus, the mean utility model is now hypothesized to be

$$(1.0) \quad U_j = \alpha [I_j + \sum_{i \neq j} I_i / D_{ij}] / D_j,$$

where:  $D_{ij}$  is the perceived cost of travel between centers  $i$  and  $j$ ,  
 $D_j$  is the perceived cost of travel from origin (home) to site  $j$ ,  
 and  $I_j + \sum_{i \neq j} I_i / D_{ij}$  is the net drawing power of center  $j$ .

The general model (1.0) can be operationalized with a number of empirical measures. Intrinsic drawing power,  $I_j$ , typically has been approximated by a size proxy such as the square footage of selling space in a center while cost,  $D_j$ , has been measured by driving time or geographical distance. The impetus for using these measures has been, of course, that they are obtainable on an objective, non-obtrusive basis at less cost than other measures which require field surveys. It should be clear, however, that other measures could be used to operationalize (1.0). For example, measures of consumers' perceptions and preferences through non-metric multidimensional scaling or conjoint measurement could possibly be used to find utility scales that would provide a more epistemic isomorphism between the theory and its empirical operants. Furthermore, the parameter estimation scheme which we develop can in theory accommodate any composition of measures making up  $I_j$  or  $D_j$ . For further exposition, however, we operationalize (1.0) using the more traditional measures as this allows a more direct comparison with current models.

Thus, using square footage ( $S$ ) as a measure of  $I_j$  and distance ( $d$ ) as a measure of  $D_{ij}$  and  $D_j$  yields the operational estimation

$$(2.0) \quad U_j = \alpha [S_j + \sum_{i \neq j} S_i / d_{ij}^{\lambda_2}] / d_j^{\lambda_1}.$$

Using the more traditional measures makes it easier to see that this formulation is a generalization of the two-body gravitational model of consumer shopping behavior. For example, suppose  $\lambda_2$  is large. This suggests a high disutility or cost in traveling between centers (assuming  $d_{ij} > 1$ ), and in such a case  $U_j$  reduces approximately to the two-body formulation (exactly as  $d_{ij}$  or  $\lambda_2 \rightarrow \infty$ ). Empirically, the  $n$ -body model could be used as a test statistic for the two-body model in the sense that large estimates of  $\lambda_2$  would tend to support the two-body formulation and small values would tend to refute it (assuming that  $S_j$  and  $d_j$  are effective proxies for utility and disutility).

The chief advantage of generalizing the two-body model to a multiple center one is that the  $n$ -body model must have at least equivalent predictive power and yet it allows for cases where consumers' perceptions of the utility of a center cannot realistically be assumed as independent of other centers. The primary disadvantage is the incorporation of the additional parameter  $\lambda_2$  and attendant parameter estimation complications.

### The Study

Data for this study was obtained from the Peoria, Illinois trading area. The central business district and five shopping centers were specifically focused on with shoppers selected systematically from each center. The square footage estimated for the CBD was 2,100,000 while corresponding estimates

ranged from 100,000 to 750,000 square feet for the centers. The sample size for each center varied from 350 to 625 respondents. By combining selected census tracts, a total of 25 residence areas were defined. Spatial disutility was approximated by Euclidian distance (units in miles). The basic responses modeled were clothing intentions data, e.g., "I came here to shop for clothing" obtained via personal interview.

Estimates of  $\lambda_1$  and  $\lambda_2$  were obtained in this survey by applying the maximum likelihood technique to the probability model

$$(3.0) \quad P_j = U_j / \sum_j U_j,$$

where  $U_{ij}$  is given by (2.0). Operationally, the minimum of the negative of the log likelihood function was obtained by a conjugate gradient method (the FMCG subroutine in the IBM Scientific Subroutine Package). Thus the function

$$(4.0) \quad L = -\sum_i \sum_j N_{ij} \log_e [P_{ij}(\lambda_1, \lambda_2)].$$

was minimized with respect to  $\lambda_1$  and  $\lambda_2$  where  $N_{ij}$  represents the observed number of shoppers from area  $i$  in center  $j$  and  $P_{ij}(\lambda_1, \lambda_2)$  is the theoretical probability function given in (3.0).

### The Results

Table 1 depicts observed probabilities and those estimated by the two-body and n-body models. The natural log of square footage was used in the n-body model as a proxy for utility. While there are substantial deviations from the observed probabilities for both models, the n-body model clearly fits much better. For example, the sum of squared deviations from observed probabilities (SSD) was 7.442 for the two-body model and 2.045 for the n-body model. Obviously, the disproportionate size of the CBD distorts true utility. Thus the SSD differential is due both to the additional parameter of the n-body model plus the log transformation of utility.

Table 2 is a comparison of the two-body and n-body models where log square footage is used for both models. The results here are quite similar, though the n-body model exhibits a slight superiority by the SSD criterion -- SSD = 2.045 versus 2.790 for the two-body model.

The estimates of the parameter  $\lambda_1$  were quite similar for both models. When the log transformation was used,  $\lambda_1$  was 1.552 for the two-body model and 1.514 for the n-body model. The estimate for the two-body model without the log transformation was 2.272. This is consistent with expectations since the log transformation attenuates the numerator and the smaller value of  $\lambda_2$  reduces the value of the denominator.

The estimates obtained for  $\lambda_2$  were unexpected. According to the rationale developed earlier, the added contribution to utility by a center increases as

TABLE 1  
EMPIRICAL, TWO-BODY, AND N-BODY PROBABILITIES: N-BODY  
SQUARE FOOTAGE LOG TRANSFORMED.

Neighborhood	Center					
	1	2	3	4	5	6
1.	.559 .227 <sup>a</sup> (.082)	.132 .242 (.313)	.087 .091 (.061)	.025 .000 (.074)	.043 .030 (.059)	.154 .409 (.411)
2.	.774 .207 (.200)	.038 .103 (.170)	.062 .138 (.079)	.014 .000 (.075)	.034 .052 (.083)	.079 .500 (.393)
3.	.959 .600 (.655)	.005 .029 (.070)	.017 .057 (.061)	.003 .000 (.048)	.010 .171 (.073)	.006 .143 (.094)
4.	.925 .526 (.533)	.007 .105 (.075)	.032 .105 (.097)	.006 .053 (.073)	.023 .105 (.129)	.007 .105 (.093)
5.	.767 .476 (.261)	.015 .048 (.094)	.104 .286 (.171)	.017 .000 (.117)	.084 .143 (.257)	.014 .048 (.101)
6.	.870 .240 (.402)	.011 .080 (.092)	.058 .080 (.134)	.008 .080 (.081)	.043 .400 (.189)	.010 .120 (.103)
7.	.775 .347 (.244)	.020 .030 (.118)	.094 .238 (.143)	.011 .297 (.071)	.062 .257 (.179)	.039 .099 (.245)
8.	.361 .200 (.071)	.013 .000 (.047)	.212 .292 (.186)	.011 .031 (.042)	.393 .462 (.614)	.010 .015 (.041)
9.	.639 .256 (.175)	.016 .000 (.080)	.151 .308 (.193)	.015 .051 (.082)	.167 .231 (.393)	.014 .154 (.077)
10.	.261 .094 (.058)	.015 .031 (.059)	.552 .516 (.512)	.010 .016 (.046)	.153 .297 (.281)	.010 .047 (.044)
11.	.276 .147 (.053)	.012 .029 (.040)	.310 .324 (.254)	.012 .000 (.045)	.382 .471 (.574)	.009 .029 (.035)
12.	.275 .193 (.061)	.016 .011 (.062)	.547 .409 (.507)	.017 .000 (.072)	.135 .330 (.250)	.011 .057 (.049)
13.	.437 .333 (.096)	.019 .000 (.074)	.312 .111 (.302)	.025 .056 (.106)	.193 .389 (.355)	.015 .111 (.067)
14.	.480 .143 (.106)	.024 .000 (.093)	.297 .429 (.291)	.038 .000 (.161)	.142 .286 (.267)	.018 .143 (.082)
15.	.440 .000 (.096)	.025 .000 (.093)	.343 1.000 (.328)	.036 .000 (.150)	.138 .000 (.256)	.018 .000 (.078)
16.	.427 .167 (.092)	.035 .000 (.127)	.372 .667 (.348)	.034 .000 (.138)	.112 .000 (.207)	.021 .167 (.089)
17.	.543 .097 (.130)	.018 .000 (.0756)	.205 .290 (.223)	.021 .032 (.098)	.199 .484 (.401)	.015 .097 (.072)
18.	.612 .180 (.122)	.118 .195 (.387)	.145 .172 (.136)	.045 .070 (.172)	.064 .094 (.116)	.016 .289 (.067)
19.	.531 .200 (.082)	.044 .100 (.115)	.139 .200 (.099)	.188 .500 (.508)	.064 .000 (.089)	.036 .000 (.107)
20.	.849 .000 (.319)	.017 .100 (.118)	.056 .200 (.106)	.020 .100 (.153)	.034 .400 (.120)	.024 .200 (.183)
21.	.719 .217 (.171)	.030 .044 (.125)	.095 .217 (.109)	.067 .370 (.301)	.052 .076 (.115)	.038 .076 (.180)
22.	.498 .032 (.067)	.035 .060 (.082)	.111 .097 (.071)	.270 .774 (.629)	.054 .065 (.066)	.032 .032 (.085)
23.	.418 .085 (.048)	.029 .000 (.058)	.101 .014 (.055)	.378 .845 (.729)	.049 .028 (.051)	.026 .028 (.059)
24.	.306 .051 (.029)	.020 .000 (.033)	.072 .051 (.032)	.549 .898 (.841)	.035 .000 (.030)	.018 .000 (.035)
25.	.718 .286 (.202)	.021 .000 (.105)	.124 .333 (.166)	.038 .048 (.207)	.079 .238 (.201)	.020 .095 (.118)

<sup>a</sup>Table entries are observed probabilities (centered), two-body probabilities and n-body probabilities (in parentheses).

TABLE 2  
 EMPIRICAL, TWO-BODY, AND N-BODY PROBABILITIES:  
 SQUARE FOOTAGE LOG TRANSFORMED FOR BOTH MODELS.

Neighborhood	Center					
	1	2	3	4	5	6
1.	.148 .227 <sup>a</sup> (.082)	.302 .242 (.313)	.063 .091 (.061)	.030 .000 (.074)	.059 .030 (.059)	.397 .409 (.411)
2.	.306 .207 (.200)	.162 .103 (.170)	.074 .138 (.079)	.029 .000 (.075)	.075 .052 (.083)	.354 .500 (.393)
3.	.740 .600 (.655)	.060 .029 (.070)	.049 .057 (.061)	.016 .000 (.048)	.055 .171 (.073)	.081 .143 (.094)
4.	.648 .526 (.533)	.067 .105 (.075)	.077 .105 (.097)	.025 .053 (.073)	.098 .105 (.129)	.085 .105 (.093)
5.	.395 .476 (.261)	.096 .048 (.094)	.150 .286 (.171)	.043 .000 (.117)	.211 .143 (.257)	.105 .048 (.101)
6.	.532 .240 (.402)	.086 .080 (.092)	.110 .080 (.134)	.028 .080 (.081)	.146 .400 (.189)	.098 .120 (.103)
7.	.361 .347 (.244)	.114 .030 (.118)	.124 .238 (.147)	.027 .297 (.071)	.148 .257 (.179)	.227 .099 (.245)
8.	.142 .200 (.071)	.060 .000 (.047)	.188 .292 (.186)	.020 .031 (.042)	.537 .462 (.614)	.054 .015 (.041)
9.	.290 .256 (.175)	.087 .000 (.080)	.176 .308 (.193)	.033 .051 (.082)	.326 .231 (.393)	.088 .154 (.077)
10.	.118 .094 (.058)	.073 .031 (.059)	.466 .516 (.512)	.021 .016 (.046)	.264 .297 (.281)	.057 .047 (.044)
11.	.111 .147 (.053)	.053 .029 (.040)	.254 .324 (.254)	.022 .000 (.045)	.514 .471 (.574)	.047 .029 (.035)
12.	.125 .193 (.061)	.077 .011 (.062)	.464 .409 (.507)	.033 .000 (.072)	.239 .330 (.250)	.064 .057 (.049)
13.	.182 .333 (.096)	.088 .000 (.074)	.284 .111 (.302)	.045 .056 (.106)	.319 .389 (.355)	.082 .111 (.067)
14.	.200 .143 (.106)	.109 .000 (.093)	.276 .429 (.291)	.066 .000 (.161)	.250 .286 (.267)	.100 .143 (.082)
15.	.184 .000 (.096)	.109 .000 (.093)	.309 1.000 (.328)	.062 .000 (.150)	.241 .000 (.256)	.096 .000 (.078)
16.	.175 .167 (.092)	.143 .000 (.127)	.323 .667 (.348)	.057 .000 (.138)	.196 .000 (.207)	.107 .167 (.089)
17.	.231 .097 (.130)	.087 .000 (.756)	.210 .290 (.223)	.041 .032 (.098)	.346 .484 (.401)	.086 .097 (.072)
18.	.221 .180 (.122)	.381 .195 (.387)	.135 .172 (.136)	.068 .070 (.172)	.144 .094 (.116)	.080 .289 (.067)
19.	.198 .200 (.082)	.164 .100 (.115)	.131 .200 (.099)	.232 .500 (.508)	.116 .000 (.089)	.159 .000 (.107)
20.	.459 .000 (.319)	.114 .100 (.118)	.095 .200 (.106)	.054 .100 (.153)	.103 .400 (.120)	.175 .200 (.183)
21.	.309 .217 (.171)	.142 .044 (.125)	.114 .217 (.109)	.116 .370 (.301)	.117 .076 (.115)	.203 .076 (.180)
22.	.189 .032 (.067)	.137 .000 (.082)	.110 .097 (.071)	.318 .774 (.629)	.100 .065 (.066)	.147 .032 (.085)
23.	.160 .085 (.048)	.115 .000 (.058)	.099 .014 (.055)	.416 .845 (.729)	.090 .028 (.051)	.121 .028 (.059)
24.	.121 .051 (.029)	.083 .000 (.033)	.074 .051 (.032)	.565 .898 (.841)	.068 .000 (.030)	.089 .000 (.035)
25.	.338 .286 (.202)	.115 .000 (.105)	.157 .333 (.166)	.078 .048 (.207)	.182 .238 (.201)	.131 .095 (.118)

<sup>a</sup>Table entries are observed probabilities (centered), two-body probabilities and n-body probabilities (in parentheses).

distance decreases (since  $1/d_{ij}^{\lambda_2}$  increases as  $d_{ij}$  decreases). However, our estimates of  $\lambda_2$  were negative. For the log n-body model  $\lambda_2$  was estimated to be -1.803. The most obvious explanation for this result is that, in effect, center interaction reduces utility. Indeed, we posited that  $U(A+B) > U(A) + U(B)$ , where, for example,  $U(A)$  is the utility of center A. Instead, this suggests the possibility that, essentially,  $U(A+B) < U(A) + U(B)$ . Strictly speaking, nearby centers must add utility under our formulation, but the negative exponent implies that more utility is added by distant centers -- thus, except for an additive constant, the implication is that nearby centers subtract more utility than distant centers.

While the interaction effect explanation is plausible, the role of the CBD in our estimates perhaps offers the more likely interpretation. It is likely that even under the log transformation the utility of the CBD was overstated. In attempting to compensate for this, the role of  $\lambda_2$  was to further minimize the effect of the CBD. Since the CBD is near the centroid of the shopping area, the distance to other centers is, on the average, smaller than the analogous statistic for other centers. Thus, a negative exponent  $\lambda_2$  makes distance a positive utility factor and the effect is to increase the utility of outlying shopping centers relative to that of the CBD.

Some intuitive modifications may possibly correct the overstated utility of the CBD. For example, the hours of operation were much shorter than those in shopping centers. Thus a likely utility proxy is square footage times hours per day of customer access; i.e., square footage shopping hours. Also, the ratio of parking space to square footage is a possible utility attenuating factor. Taken together, these modifications may provide a better estimate of utility and suggest the role of  $\lambda_2$  in quantifying center interdependence.

#### Model Extensions

A more comprehensive three parameter model is presently under development. The probability equation in this model follows from the Luce choice axiom. However, utility is cast in the form

$$\begin{aligned}
 (4.0) \quad U_j &= \log_e(S_j) [\pi_{i \neq j} \exp(-\alpha \log_e(S_i)) / d_{ij}^{\lambda_2}] / d_j^{\lambda_1} \\
 &= \log_e(S_j) [1 / \pi_{i \neq j} S_i^{\alpha/d_{ij}^{\lambda_2}}] / d_j^{\lambda_1}
 \end{aligned}$$

The rationale for this model is that each multiplicative term  $1/\pi_{i \neq j} S_i^{\alpha/d_{ij}^{\lambda_2}}$  either attenuates or magnifies the utility of center  $j$  depending on the value of  $\alpha$ . That is,  $\alpha < 0$  gives a value greater than one for this term while  $\alpha > 0$  yields a value less than one.  $\alpha = 0$  is consistent with center independence since the multiplicative term reduces to one in this case. Increasing distance,  $d_{ij}$ , and  $\lambda_2$  diminishes the center interaction effect.



## FOOTNOTES

1. Data for this study were provided by the Urban Affairs Institute, Bradley University.
2. Assistant Professors of Marketing.
3. Associate Professor of Marketing and Editor, Journal of Business Research.
4. A more comprehensive examination of the problems is found in Hilliard and Reynolds (1973).

## REFERENCES

- Hilliard, J. E., and Reynolds, F. D. A multiple center gravitational model of consumer shopping behavior. Unpublished paper. University of Georgia, 1973.
- Huff, D. L. Determination of intra-urban retail trade areas. Los Angeles: Division of Research, Graduate School of Business Administration, University of California, 1962.
- Luce, R. D. Individual choice behavior. New York: John Wiley and Sons, 1959.

CONCEPTUAL AND OPERATIONAL PROBLEMS WITH MARKET SHARE  
MODELS OF CONSUMER SPATIAL BEHAVIOR

David L. Huff<sup>1</sup> and Richard R. Batsell<sup>2</sup>

Luce's choice axiom provides the theoretical foundation for most models of consumer spatial behavior. However, empirical applications indicate that the conceptual properties of the axiom are not well understood. This paper reviews these properties in terms of choice behavior in a spatial context. A recent development in parameter estimation for market-share models is also discussed. Finally, suggestions for needed research are listed.

The development of consumer spatial behavior models has been limited principally to those of the market share type. Such models have received widespread usage from academicians as well as practitioners. It is apparent from a number of these applications that the conceptual properties of this type of model are not well understood. This is evident from some of the problems to which this type of model has been applied, and from the questionable interpretations that have been given to the output.

As a consequence, the purpose of this paper is to re-examine the major behavior postulates associated with market share type models of consumer spatial behavior. In addition, some operational difficulties that are often incurred in implementing the model will be discussed. Finally, a recent development using standard econometric methods of parameter estimation for this type of model is discussed.

Basic Choice Axiom

Most models of the market share type depict consumer choice behavior in terms of Luce's basic choice axiom (Luce, 1959). The axiom is:

$$(1.0) \quad P_T(x) = v(x) / \sum_{y \in T} v(y)$$

where:

$P_T(x)$  is the probability of an individual choosing alternative  $x$  from a finite set of  $T$  alternatives;  $T$  is a subset of some universal set  $U$ ; and,  $v(x)$  is a positive real-valued function  $v$  on  $T$ .

Thus, the basic proposition of this axiom is that choice behavior is best described as a probabilistic, not an algebraic, phenomenon. When a consumer reaches a decision between alternative  $a$  and  $b$ , it is assumed that there is a probability  $P(a,b)$  that the choice will be  $a$  rather than  $b$  and this probability will generally be different from 0 and 1. The consumer is assumed to be able to evaluate the elements of  $U$  along some comparative dimension and to be able to select a finite subset of  $T$  alternatives. Furthermore, for situations in which pairwise choice discrimination is imperfect, this choice axiom implies that the ratio of the probabilities of any two choice alternatives is constant and independent of any assumptions about the structure of the set of alternatives.

In addition, the choice axiom complies with the notion of transitivity, i.e., if choice alternative  $a$  is preferred over  $b$  and  $b$  over  $c$ , then,  $a$  is preferred over  $c$ . While this preference ordering is subject to random fluctuations, it is assumed that such fluctuations are in keeping with the constant probability vector.

### A Consumer Spatial Behavior Model

An initial attempt to model consumer spatial behavior in terms of Luce's basic choice axiom was made by Huff (Huff, 1962). The positive real-valued function  $v$  of  $x$  that was used in the model is:<sup>3</sup>

$$v(j) = A_j^\gamma D_{ij}^{-\lambda}$$

where  $A_j$  = an attraction index of retail facility  $j$ ;

$D_{ij}$  = the accessibility of a retail facility  $j$  to a consumer located at  $i$ ; and

$\gamma, \lambda$  = empirically determined parameters.

The probability that a consumer located at  $i$  will choose to shop at retail facility  $j$  for a particular product or group of homogeneous products is:

$$(2.0) \quad p_{ij} = \frac{A_j^\gamma D_{ij}^{-\lambda}}{\sum_{j=1}^n A_j^\gamma D_{ij}^{-\lambda}}$$

The quotient derived by dividing  $A_j^\gamma$  by  $D_{ij}^\lambda$  is regarded as the perceived utility of retail facility  $j$  by a consumer located at  $i$ . Square footage of selling space is often used as a surrogate measure for the attraction variable  $A_j$ . Travel time, road and straight-line distance, and travel costs are measures that have been used for the accessibility variable  $D_{ij}$ . The parameter associated with a variable reflects the sensitivity of the probabilities with respect to that variable. A comparison of the values of these parameters for different types of shopping trips is particularly meaningful. Various estimation procedures have been employed to estimate parameters of non-linear models (Wilson, 1974). In general, the procedures involve specifying some initial value for one of the parameters and then calculating a value for the other using some efficient search routine. This iterative procedure is continued until the lowest sum of squared differences is reached between the actual and expected number of consumers shopping at a specified set of retail facilities.

### Problem Areas

#### Trip-type

Consumer spatial behavior as described by equation (2.0) pertains to single-purpose shopping trips. However, many shopping trips are multi-purpose in nature. These trips may involve multi-product purchase intentions in which different retail facilities are visited. Or, they may involve single-purpose shopping trips that are made in conjunction with non-shopping activities. In the case of multi-purpose trips, the proximity of a retail facility to a consumer's place of origin may not be nearly as important as the proximity of a retail facility to other retail facilities in which purchases are intended,

or, to those non-shopping activity places that the individual intends to visit. Therefore, if the model is used, it is important that the set of retail facilities that are specified as choice alternatives be those that are most likely to be associated with single-purpose shopping trips.

### Product-Type

The model of consumer spatial behavior is formulated to specify the likelihood of a consumer choosing a given retail facility to purchase a specific product, e.g., an automobile, or, a group of related products, e.g., groceries. Therefore, it is important that the attraction values that are specified for those retail facilities comprising the set of choice alternatives be in keeping with the product(s) purchase intentions of the consumer. This point is often overlooked by those employing the model. Frequently, no product or group of related products are specified except in a very aggregative way, e.g., retail goods and services. Similarly, in those cases where a specific product(s) is designated, the measure of attraction for the retail facilities is often an aggregative measure that reflects the attraction for other products as well, e.g., total square footage of a department store when only purchases of women's jewelry are being analyzed.

### Spatial Equilibrium

The essence of a market share model of consumer spatial behavior is that a consumer may share his patronage among feasible alternatives and his preference ordering will be in keeping with a constant probability vector of the form (.60, .30, .10). The model does not indicate how the market-sharing principle was established, and it provides an equilibrium solution without knowing if a consumer is, in fact, in equilibrium. Furthermore, the probability vector is derived from cross-sectional data which may not be a true indication of spatial equilibrium conditions (Golledge, 1970). It seems reasonable that for most resident populations there will be some consumers who are in the initial search phases of behavior, some with partially-formed preference patterns, and others with well-formed patronage habits.

### Choice Alternatives

Problems may also arise in using the model if the subset of choice alternatives is not defined correctly. Two types of errors can occur. First, if a choice alternative which is actually part of a consumer's set has been excluded, then although the ratio of probabilities between any pair of choice alternatives will not be affected, errors of prediction can occur. In this case, expected values (e.g., annual expenditures) for those choice alternatives which were included in the specified subset will be over-estimated.

Secondly, if a choice alternative which is not actually in a consumer's set of choice alternatives is included, then expected value errors will depend on model specification. For example, if an individual does not consider some particular choice alternative then there is probably some characteristic of the choice alternative that distinguishes it from those alternatives which are considered. If this characteristic is reflected as a variable in the model specification, then the model will be able to discriminate effectively between choice alternatives. Consequently, differences between the observed and expected values will not be the result of defining the set improperly. Conversely, if the differentiating variable is not specified in the model, then the effect of not defining the choice alternatives properly can result in marked differences between observed and expected values.

### Group Behavior

Another potential problem arises when the model is used to examine the choice behavior of groups of consumers. The original choice axiom applies specifically to individual choice behavior. However, most applications of the model are oriented toward the analysis of group behavior. Typically, a geographic area is divided into sub-areas and the proportion of times consumers patronize places is averaged for each of these areas.<sup>4</sup> The assumption is that the smaller the geographic area the more homogeneous are the residents and thus the more similar their spatial behavior. The parameters of the model are estimated for each sub-area and then averaged across all sub-areas for prediction purposes. If the parameter estimates for each of the sub-areas vary considerably, then the average of these values will be misleading and result in incorrect predictions.

### Choice Determinants

A consumer's perceived utility of a retail facility is unquestionably determined by a myriad of factors. However, equation (2.0) specifies only two variables even though these variables are regarded as surrogates for a number of different variables. While other variables could be included in the model, there is some difficulty in estimating the parameters associated with such variables. Furthermore, the techniques traditionally used in estimating the parameters do not assure that a global maxima or minima will be obtained nor are the statistical properties of the estimates known. Due to the non-linear formulation of the model most researchers believed that the parameters could not be estimated by standard econometric methods. However, recently Nakanishi and Cooper demonstrated that the parameters can be estimated using least squares (Nakanishi and Cooper, 1974). A generalization of equation (2.0) with respect to consumer spatial behavior is as follows:

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

where:

$\pi_{ij}$  = the probability that a consumer located at  $i$  will choose a retail facility  $j$ ;

$X_{kij}$  = the  $k^{\text{th}}$  of  $q$  variables describing a retail facility  $j$  in terms of a consumer located at  $i$ ;

$\beta_k$  = the sensitivity of  $\pi_{ij}$  with respect to variable  $k$ ; and

$m$  = the number of retail facilities.

After making a log transformation and rearranging terms, the model is of the form:

$$y_{ij} = \sum_{k=1}^q \beta_k Z_{kij} + E_{ij}$$

where:

$$y_{ij} = \log \left( \frac{\pi_{ij}}{\bar{\pi}_i} \right) \text{ and } \bar{\pi}_i \text{ is the geometric mean of the } m \text{ different probabilities corresponding to each of the consumers;}$$

$$z_{kij} = \log \left( \frac{X_{kij}}{\bar{X}_{ki}} \right) \text{ and } \bar{X}_{ki} \text{ is the geometric mean of the } m \text{ different values of } X_{kij} \text{ describing the retail facilities in terms of variable } k \text{ for a consumer located at } i; \text{ and}$$

$E_{ij}$  = the difference between the model estimates of  $y_{ij}$  and the actual values of  $y_{ij}$ .

Thus the original two-variable model can be extended to  $q$  variables and the parameters associated with these variables can be estimated using least-squares techniques. For example, if the effects of price, distance and size on consumer spatial behavior are to be assessed, the model would be expressed as follows:

$$\pi_{ij} = \frac{S_{ij}^{\beta_S} \cdot D_{ij}^{\beta_D} \cdot L_{ij}^{\beta_L}}{\sum_{j=1}^m S_{ij}^{\beta_S} \cdot D_{ij}^{\beta_D} \cdot L_{ij}^{\beta_L}}$$

where:

$\pi_{ij}$  = the probability that a consumer in sub-area  $i$  will choose location  $j$ ;

$S_{ij}$  = the square footage of selling space in retail location  $j$ ;

$\beta_S$  = the sensitivity of  $\pi_{ij}$  with respect to store size;

$D_{ij}$  = the distance between sub-area  $i$  and location  $j$ ;

$\beta_D$  = the sensitivity of  $\pi_{ij}$  with respect to distance;

$L_{ij}$  = the price level for a given consumer at  $i$  with respect to retail location  $j$ ;

$\beta_L$  = the sensitivity of  $\pi_{ij}$  with respect to the price level;  
and

$m$  = the number of retail locations.

The transformed model would be:

$$Y_{ij} = \beta_S \log \left( \frac{S_{ij}}{\bar{S}_i} \right) + \beta_D \log \left( \frac{D_{ij}}{\bar{D}_i} \right) + \beta_L \log \left( \frac{L_{ij}}{\bar{L}_i} \right)$$

where:

$\beta_S$ ,  $\beta_D$ ,  $\beta_L$ ,  $S_{ij}$ ,  $D_{ij}$ , and  $L_{ij}$  are defined as before;

$Y_{ij} = \log \left( \frac{\pi_{ij}}{\bar{\pi}_i} \right)$  and  $\bar{\pi}_i$  is defined as the geometric mean of the  $\underline{m}$  different probabilities associated with a consumer at  $\underline{i}$ ;

$\bar{S}_i$  = the geometric mean of the  $\underline{m}$  different retail square footage values;

$\bar{D}_i$  = the geometric mean of the distances between the location of a consumer at  $\underline{i}$  and the  $\underline{m}$  different retail locations; and

$\bar{L}_i$  = the geometric mean of the price level for the  $\underline{m}$  different retail locations for a consumer located at  $\underline{i}$ .

There are distinct advantages to this approach.

1. Additional variables with potential explanatory power can easily be added to the model and the associated sensitivity parameters estimated.
2. Using alternate forms of full and restricted linear models, the importance of each variable can be tested.
3. The solution for the least-squares estimates of the sensitivity parameters assures a global minimum.
4. If the appropriate least-squares technique is used, the parameter estimates will be unbiased and of minimum variance.
5. Computer algorithms for least-squares estimation are readily available.

However, it should be noted that there still exist certain limitations.

1. If any one of the proportions are zero, the geometric mean will be zero and that particular proportion cannot be transformed.
2. The traditional assumptions necessary for tests of significance of alternate linear models apply.

### Conclusions

Considerably more empirical research of consumer spatial behavior is needed in order to assess the validity of the Luce choice axiom. Such research should be structured in keeping with the conceptual properties underlying the choice axiom. Examples of needed research include:

1. What particular products or groups of related products involve single-purpose shopping trips?
2. How many retail outlets normally comprise a consumer's set of choice alternatives with respect to the purchase of a given product or group of related products?
3. Does the number of outlets comprising the set of choice alternatives vary not only by product but also in terms of the socio-economic characteristics

of consumers?

4. How stable are consumer patronage patterns over time with respect to different product purchases? Do these patterns vary among different types of consumers?
5. What are the perceived attributes as well as the objective equivalences that determine the utility of a retail facility to a consumer?<sup>5</sup>
6. How similar are these attributes, and their relative importance, among differing types of consumers?

#### FOOTNOTES

1. Professor of Resources and Geography, Graduate School of Business, The University of Texas at Austin.
2. Research Associate, Graduate School of Business, The University of Texas at Austin.
3. The initial model specified only one parameter, an exponent on distance. However, subsequent refinements included an additional exponent on the attraction index variable.
4. An excellent study exemplifying this approach was conducted by Haines, et. al. (Haines, Simon and Alexis, 1972).
5. Research pertaining to the spatial aspects of this question has been done by Briggs (Briggs, 1969), Burnett (Burnett, 1973), and MacKay (MacKay, Olshavsky, and Sentell, 1975).

#### REFERENCES

- Briggs, R. The scaling of preferences for spatial locations: An example using shopping centers. Unpublished M.S. thesis, The Ohio State University, 1969.
- Burnett, P. The dimensions of alternatives in spatial choice processes. Geographical Analysis, 1973, 5, 181-204.
- Golledge, R. G. Some equilibrium models of consumer behavior. Economic Geography, 1970, 46, 417-423.
- Haines, G. H., Simon, L. S., & Alexis, M. An analysis of central city neighborhood food trading areas. Journal of Regional Science, 1972, 12, 95-105.
- Huff, D. L. A probabilistic analysis of consumer spatial behavior. In W. S. Decker (Ed.), Emerging Concepts in Marketing. Chicago: American Marketing Association, 1962(b).
- Huff, D. L. The delineation of a national system of planning regions on the basis of urban spheres of influence. Regional Studies, 1973, 7, 323-329.



Luce, R. D. Individual Choice Behavior. New York: John Wiley & Sons, 1959.

Mackay, D. B., Olshavsky, R. W. & Sentell, G. Cognitive maps and spatial behavior of consumers. Geographical Analysis, in press.

Nakanishi, M. & Cooper, L. G. Parameter estimation for a multiplicative competitive interaction model--least squares approach. Journal of Marketing Research, 1974, 11, 303-311.

RETAIL MARKET AREA SHAPE AND STRUCTURE:  
PROBLEMS AND PROSPECTS

J. Barry Mason<sup>1</sup>  
The University of Alabama

The keys to understanding the shape of market areas are found in central place literature. Our concepts of market area structure are anchored in the refinements of the traditional gravity model. Central place theory provides the key assumptions underlying the structure of market areas. Thus, structure emerges as a subset of shape. However, we must ask ourselves the purpose we want predictive mechanisms to serve. If our needs are satisfied by gross decision parameters at the level of the firm, then perhaps our models of shape are satisfactory and our structural requirements are not overly monumental. If, however, we are seeking the basics for a theory of consumer demand at the micro-analytic level, then our structural models particularly are woefully inadequate.

Marketers frequently use various types of spatial analyses in market area planning without understanding the underlying bases of the implicit spatial and temporal relationships in the paradigms used. Most students of consumer behavior are familiar with the contributions of Reilly, Converse, and perhaps one or two other individuals (Schwartz, 1963; Don and Ruth Mulvihill, 1970; Revzan, 1968) such as Huff but their knowledge for the most part does not go beyond this. Perhaps a reason for relative lack of interest is that many researchers do not believe that these paradigms offer the field of marketing explicit means for explaining or predicting the spatial behavior of consumers (Moore and Mayer, 1966). Thus, marketers have turned to a variety of other disciplines in developing useful spatial and temporal constructs for market analysis. Primary contributions have been made by geographers (Berry and Horton, 1970), economic historians (Fite and Reese, 1965), land economists (Hoover, 1948), regional scientists (Isard, 1956), and others.

The explanations offered by these individuals are quite varied. For example, one group of researchers focused on the external relationships of a city to surrounding regions. This thrust includes contributions by such writers as Von Thunen (1966), Weber (1929), Losch (1959), and Greenhut (1952). These writers have been primarily concerned with industrial site location as well as land utilization and spatial supply and demand relationships. Another group of writers, including Fetter (1924), Christaller (1966), and Isard (1960), focused on the intra- and inter-relationships of the city and surrounding regions. Still other contributions have been made by such groups as the human ecologists, whose work is exemplified by R. D. McKenzie and R. F. Park (Mason, 1973), and by the social physicists (Warntz, 1950).

Of particular interest in this paper, however, is a specific focus on the various dimensions of retail market areas. A series of generalizations have emerged on the shape and structure of market areas but gaps still exist in our knowledge. Thus, the specific foci of this paper are (1) to review existing generalizations on the shape and structure of retail markets; (2) to review existing

generalizations on consumer behavior and market organization as related to the shape and structure of markets; and (3) to identify existing gaps in our knowledge of the spatial and temporal dimensions of consumer and market behavior and to offer suggestions for further research.

### The Shape of Market Areas

Numerous writers have contributed to the literature on the shape of retail markets. The major theoretical breakthroughs, however, were by Walter Christaller and August Losch (Baskin, 1957), both of whom base their theories on the behavior of retailers and consumers over time and in space.

Christaller's theory begins with the smallest unit of human settlement which provides goods and services for a surrounding region. He termed these central functions and the surrounding region a central place, which he stated would be located at the center of minimum aggregate travel. He further assumes identical consumers, distributed with uniform densities over an unbounded plain with equal accessibility from all points of the surrounding region, and uniform purchasing power. He then determined that a hexagonal market area would require the least average distance for movement to the centers by consumers. He then assumes a hierarchy of settlements with a distinct order of magnitude (size). Each order contains a fixed number,  $k$ , of settlements for each settlement of the order above. For instance, in Christaller's model  $k$  is assumed to be seven; thus, for each city there will be six surrounding towns (in addition to the town contained within the city) in its direct sphere of influence and for each town six surrounding villages. The assumption of a constant  $k$  restricts the number of consumers in a market area. Also, the larger the center the more extensive is the market area and the greater the specialization of services. Lastly, to assign different activities to different markets, Christaller introduces the concept of the "range of a good" which establishes the level below which an insufficient number of consumers exists to support a specified type of activity (Bunge, 1973). Christaller's resulting market areas are shown in Figure 1.

August Losch also starts with the smallest unit of settlement, but proceeds differently from Christaller. He assumes that the cost of transportation increases with distance. Thus, price is higher at the edges of a market and the demand is lower. This result yields a demand cone, as shown in Figure 2, where  $PQ$  equals the quantity demanded at the market center. Demand is assumed to decrease along  $PF$  as distance increases from  $Q$  to  $F$ . He then computes the volume of the demand cone as a measure of total demand and shows that when a plain is filled with market areas and the demand curve  $PF$  is assumed to be a straight line, a hexagonal based cone as shown in Figure 3 will yield more volume than any other configuration (Bunge, p. 136-137).

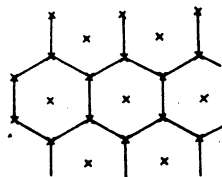


Figure 1.

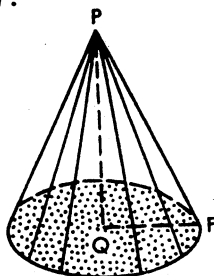


Figure 2.

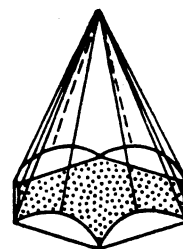


Figure 3.

Source: William Bunge, Theoretical Geography, Lund Studies in Geography, Series C, General and Mathematical Geography, No. 1. The Royal University of Lund, Sweden (C.W.K. Gleerup, 1973) pp. 136-137.

In a manner comparable to Christaller, Losch assumes uniform distribution of population and hexagonal market areas. However, his system does not have the rigidity of Christaller in that he does not assume, except as a special case, a fixed  $k$ . Thus a larger number of market areas may exist. Further, larger settlements do not necessarily contain all of the activities of smaller settlements or the same activities in communities of the same size.

The noted geographer Brian Berry (1967, p. 72) highlights the Christaller and Losch models as follows:

Both Christaller and Losch agree that the triangular arrangement of production sites or retail stores, and hexagonal market areas, represents an optimum for a single good, under the assumption of uniform densities on an unbounded plain, with equal access in all directions. Losch provides explicit proof of this. In the solution, the location of individual firms is as advantageous as possible, every consumer receives service, abnormal profits disappear, market areas are as small as possible, and the boundaries of market areas are points of consumer indifference....

Christaller's formulation appears most relevant for understanding the geography of retail and service business, whereas that of Losch provides a framework for analyzing the spatial distribution of market-oriented manufacturing. Christaller's agglomerative requirement, for example, is compatible with the idea of travel-cost minimization by rational consumers on multi-purpose trips, a condition not satisfied by Losch.

An additional refinement of Christaller's work was made by Brian Berry and William Garrison (1958). They removed the limiting assumption of uniform purchasing power which is essential to the arrangement of hexagonal market areas. Their theory was reformulated in terms of threshold and range. Threshold was defined as the minimum purchasing power necessary to support the supply of a good from a center. Christaller earlier defined the range as the maximum distance a consumer would be willing to travel to purchase a good offered at a given place. This limit is of course determined by competition among the centers supplying the same good, as well as other factors.

Berry and Garrison, however, have been criticized for equating threshold functions with the population of central places rather than with that of their market areas (Haggert and Gunawardena). Indeed, much of the criticism of central place theory focuses on its overly narrow concept of demand. Further, various studies have shown that the assumptions of the model are affected by such factors as varying population density, physical isolating mechanisms, intervening opportunities, community size, shopping habits, and differing socio-economic profiles of consumers (Mason and Moore, 1970). Lastly, it has been established empirically that consumers do not necessarily patronize the nearest center, as assumed by central place theory.

Further, the literature on central place theory does not make a major distinction between interurban and suburban markets. One can argue that theoretically the market boundaries established by central place theory are arbitrary. However, consumer patronage in rural areas tends to yield patterns compatible with the theory. As Peter Simons (1973-74, p.72) says, "Results indicate that distance is an important influence on consumer choice in rural areas and that an assumption of rational spatial action may have some relevance in these situations. So it seems reasonable to suggest that, in the interurban case, the shape of the

profile derived from central place theory may be an adequate first approximation of the structure of retail market areas."

Finally, Peter Scott (1970, p. 16) has commented as follows in summarizing the criticisms of central place theory as it exists today:

Undoubtedly the most serious limitations of central place theory...stem not so much from its treatment of demand, important though these limitations are, as from its concept of production. Since in retailing the decision-making unit is the firm, and most retail firms sell a variety of goods the range of a good is affected by the product mix and the pricing policy of individual establishments. These considerations are in turn influenced by the organisational structure of firms and the extent to which they enjoy internal economies of scale....But because the central place model is static, deterministic, and retrospective it cannot be used to predict the functions of shopping centres given the differential growth of multi-product firms and the organisational behavior of retailing. Nor can central place theory shed light on the internal arrangements of shops within centers, which necessarily invoke considerations of site, rent, and external economies of scale.

In spite of these criticisms, however, Berry and Pred (1961) conclude that empirical evidence from different parts of the world supports in general terms the concept of a hierarchial structure of retail outlets and centers (also, Olson, 1965).

Based on an overall evaluation of the various studies relative to the shape of trading areas, two well-known marketing geographers, William Applebaum and Saul Cohen (1961), have concluded that the shape of trading areas is not circular in the most general terms, but is elliptical, with the longer axis tending away from the CBD (Applebaum and Cohen, 1961). Simons (1973-74, p.68) states that "if a set of these irregular ellipses is converted to irregular hexagons, the result is a replica of that developed by Isard when he adjusted Losch's regular hexagons."

However, the assumption in central place theory that a central place will have complete and uniform penetration within the defined market area is unrealistic and one must turn to other sources of information for generalizations on consumer behavior within defined market areas. Thus, a focus on the internal structure of markets is necessary for a more nearly complete portrait of retail markets.

### The Structure of Market Areas

The gravity model in one form or another is the basis for most generalizations on the internal structure of markets. Broadly speaking, the gravity concept postulates that an attracting force of interaction between two areas of human activity (perhaps two cities) is created by two population masses and that a friction of movement is created by the spatial separation over which the interaction must occur (Carrothers, 1956). Interaction between two areas of population is a direct function of the size of the masses and an inverse function of the distance between them.

The earliest formulation of this concept was by H. C. Carey (1858). E. G. Ravenstein (1885) presented empirical evidence that migratory movement is to cities of large population and that the volume of movement decreases with distance.

George Zipf (1946) developed a different gravitational concept to determine the "force" of interaction between two concentrations of population. He stated that the number, sizes, and locations of communities theoretically depend on the minimization of the work of transporting mass over distance. Therefore, there are equilibria between opposing forces of unification and diversification. As a result, the interchange of items between communities  $P_1$  and  $P_2$  will be inversely proportionate to their intervening easiest transportation distance  $D$ . Given knowledge of  $D$ , the interchange of the masses (given knowledge of price-quantity relationships) will be directly proportionate to  $\frac{P_1 P_2}{D}$  for any two cities in the economy.

Samuel A. Stouffer (1940) modified the gravity flow concept in attempting to measure the mobility of population. The concept, which was empirically studied, assumes no necessary relationship between mobility and distance; rather, the concept of intervening opportunities was introduced. The contention was that "the number of persons going a given distance is directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities."

Traditional gravity formulations, as is evident from above, however, are inadequate to explain the complexity of factors that determine the economic potential for an intra-city population to support a proposed retail establishment. For example, the gravity concept is based on a notion of *ceteris paribus*. This restriction allows a concentration only on the two basic variables of distance and mass, and the factors which can be encompassed in these variables.<sup>3</sup> This two-dimensional nature thus forces the consideration of the problem in an unrealistic frame.

These models are also designed to account only for the behavior of large groups of people.<sup>4</sup> As David Huff (1962[a], p. 64) has stated of the gravity model, "It rests on the assumption that group behavior is predictable on the basis of mathematical probability because the idiosyncracies of any one individual or small group tend to be cancelled out." Since disaggregation is at best difficult, making predictions requires that the parameters of spatial interaction must be assumed to be the same for all people (Isard, 1960). The only explicit variations that have so far been given recognition in the literature are those associated with differences generated by different trip types.

When the attempt is made to employ the model as more than a descriptive tool, even more difficult problems are encountered. (Neidercorn and Bechdolt, 1969). Specifically, a lack of theory exists to explain the workings of the model. The results obtained do not explain why observed regularities occur as they do; consequently, one is at a loss when discrepancies occur that cannot be accounted for within the model parameters.<sup>5</sup> Finally, examination of choice processes in retail site selection also is not possible, especially in the earlier models which relied wholly on secondary data and essentially administrative markets (Mason and Moore, 1969). Donald Thompson (1966, p. 6) says of the early models that "Such laws must be regarded as little more than historical 'accidents' in absence of tracing out a theoretical connection between their empirically determined weights and exponents and the corresponding behavioral variables on which they rest....Using such models for forecasting or planning is potentially disastrous in absence of knowledge of the true underlying variables." This is particularly true if the forecast doesn't work, and points out the ever present epistemic gap between concepts and operations.

In the early 1960's an attempt was made by Huff to bridge the gap between consumer behavior and the gravity formulation. Huff's model using Luce's choice axiom as the theoretical foundation (Huff, 1963 and 1964; Huff, Huff, Gambini

and Jenks, 1968) assumed the following: consumers (1) isolate a subset of alternative shopping center choices from a much larger set consisting of all possible alternatives; (2) calculate a positive measure of utility for each of these perceived alternatives; (3) distribute their retail patronage spatially in probabilistic fashion. (Huff, 1962[b]). The two key operational (empirical) variables in his model are floor space and travel time. One of its primary contributions is the ability to distribute expenditures among more than two centers by the use of probability indifference curves. The model, however, lacks the richness Huff himself called for in two excellent articles written prior to the publication of the model (1960 and 1961).

Louis Bucklin (1967) raises also the caution flag on the use of travel time in this type of model. His research revealed that travel time to different facilities varies widely by time of day and day of week. He also highlighted the influence of socio-economic factors on consumer travel and the different market areas for different types of goods. Bucklin (Oct. 1967) also gives support to the notion of a curved distance function.

Thompson (1966, p. 8) has stated of the Huff model that "the model is not fundamentally different in terms of calculation of the exponents from the model of Reilly," while Scott (1970, p. 175) also made a similar observation. Thompson (p. 9) further offers the comment that "one may question whether in the consumer choice process implicit in the definition of a subset of alternatives Huff has assumed away one of the critical variables in the choice process, i.e., the less than perfect ability of the consumer to 'see' or perceive the urban or suburban environment in all of its objective complexity."

In commenting on gravity models in general, Simons (1973-74) states as follows:

...the 'accuracy' of the model is usually a result of 'calibrating' the model to fit particular situations; that is, the distance exponent is adjusted to the data being analyzed. It is hardly surprising that this kind of procedure has led to large variations in distance exponent values....However, variations in the exponent when estimated in such a fashion are not solely attributable to the independent variable of distance. What is happening, in fact, is that distance operates as a kind of error term for all other independent variables; it is an error term for variables in the equation and for others in the area being studied. Therefore, the supposed 'accuracy' of the gravity model is largely a result of an almost classic example of circular reasoning.

Scott (1970, p. 178) has succinctly summarized the overall problems and shortcomings of the gravity model as follows:

...the retail gravity model continues to pose problems when used to predict retail requirements. The problems include inter alia the definition of the study area and its constituent zones; the availability of data; the need for analysis by type of trade, social structure, and travel mode, the choice for measures of attraction and friction, the calibration of the model given only current data, the implicit assumptions concerning consumer and entrepreneurial behavior, the conversion of sales into floor-space; and the range of technological change and institutional decision-making for which reliable forecasts cannot be made. In short, the model is undoubtedly a valuable tool for preliminary macro-analysis but it remains essentially static, descriptive, and superficial.

Thompson (1966, p. 17) correctly views the state of the literature on retail area markets when he states that "it is not possible to identify a single profile shape [or so many as to be useless], indicative of the structure of retail market areas. The indeterminant nature of the structural dimension is a result of variations in the importance of distance and in the effort consumers are willing to make to purchase goods. Existing theory in this area is either poorly developed, too restrictive, or has conceptual problems."

This paper thus far has focused essentially on the gravity model and central place theory as providing the key thrusts in research on the structure and shape, respectively, of market areas. Meaningful contributions have also been made in terms of various facts of movement theory which offer promise of breakthroughs, particularly in the structural dimensions of trading areas. However, this body of knowledge is not nearly as well developed as is the literature on gravity models. The two are obviously related, however, and movement theory does seem to offer future directions for research.

Harold Hotelling (1921) more than 50 years ago presented a theory of human movement stated in terms of heat flow. He applied his model to human migration in illustrating the historical western migration of population in the United States. Martin Beckmann (1957) has applied hydrodynamic theory to economic and population geography with meaningful results. In his work he related hydrodynamic theory to social physics and the potential of population. Paul Richards (1956) has also applied fluid theory to the movement of automobiles on roads and the friction of space which is encountered. This has yielded promising theoretical constructs on the frictions of spatial interaction.

Kinetic gas theory has also inspired relatively numerous movement theories which have been applied to ecology, to studies of contagious diseases, and to the problems of spatial distribution and spread of animals over time. Neyman and Scott (1957) provide a good literature review on spatial distribution and spread in the context of kinetic gas theory. Torsten Hagerstrand (1952) was the first geographer to develop statistical movement theory. He applied it to peoples and ideas as a foundation for much of his work on the diffusion of ideas. His findings also offer relevance for the study of consumer movement in a spatial and a temporal context. The various ideas embodied in these diverse disciplinary thrusts serve to illustrate the diversity of talent focusing on the spatial and temporal dimensions of human behavior. However, the most immediately promising research is in the micro-analytic dimensions of behavior--particularly imagery research and action space formation as these relate to interaction with the retail structure of an area.

#### Micro-Behavior: An Area of Needed Conceptual Breakthroughs

Numerous researchers have built conceptual frameworks using the costs and utilities of travel as a basis for relating consumer search propensities to retail area structure in their efforts to overcome the deficiencies of macro-models of behavior. All of these efforts, however, have been predicated on an "economic man" who objectively evaluates revenues and costs, utilities and disutilities, and who then makes a decision which will maximize these utilities (Cox, 1959). One of the more promising departures from these assumptions is imagery research as applied to the geographic distribution of retail patronage, particularly the work of Kevin Lynch. (Lynch, 1960). Despite the spatial implications of such research, the imagery concept has yet to be applied to the analysis of retail area structure. However, Thompson (1966, p. 17), following his literature review in 1966, suggested the following hypothesis: "the fundamental factor



affecting the geographic distribution of retail trade is the manner in which consumers organize their perceptions of the external environment with which they are faced." In a similar vein, John Nystuen (1967), in commenting upon efforts to develop a theory of intraurban consumer behavior, states that "existing...theory is not even adequate for considering the role of customer travel behavior in creating interaction between store types." The general conclusion of Lynch (1960) is that there is a highly subjective orientation of the individual to the urban environment.

An approach in some ways conceptually analogous to image formation is that of action space formation which has been subjected to limited empirical testing. Frank Horton and David R. Reynolds (1970), in a recent report on action space formation, prefaced their work as follows:

Isard has suggested that variations in individual space and time preferences are so great as to preclude any economic rationalization of individual travel behavior. Thus far, Isard's pessimism seems to have been justified in that deterministic economic models, with their built-in assumptions of economic rationality, have been noteworthy for their lack of success in accounting for spatial behavior, except at a highly aggregative level. It would appear appropriate, therefore, to adopt a behavioral approach that examines the formulation of the individual's action space and his resulting travel behavior as a function of his socio-economic characteristics, his cognitive images of the urban environment, and his preferences for travel.

The authors point out that all attempts to use variables useful at the aggregate level in explaining travel behavior have not been useful at the level of individual movement because "Factors that are important conditioners of mass behavior (such as employment rate, median income, etc.) are devoid of behavioral meaning at a less aggregate level. A better understanding of the household's travel behavior demands that more research be directed toward discerning fundamental processes underlying this behavior." (p. 138). Action space formation is one such process. The author concurs with their assessment and supports Thompson's assessment (1966, p. 17) that, in commenting on efforts to blend the macro and micro aspects of behavioral analysis, "running through this literature one finds the twin threads of shallow empiricism and unrelated or unapplied theory. Before further research is undertaken in the former area [macro-analysis], it would seem only logical that an attempt be made first to bridge the gap between the two [macro and micro]."

#### Needed Research

The gaps in the literature are numerous, but the following suggestions may present the way for at least a limited amount of needed additional research:

1. An effort needs to be made to determine whether the division of consumer patronage between various centers in non-urban areas is distinct or whether it tends to overlap. The key question is whether rational spatial action in terms of distance works to a greater extent in non-urban areas than in urban areas.
2. An independent method needs to be developed for estimating the distance exponents for particular situations as opposed to deriving them for each area and calibrating gravity models in this way.
3. Research needs to be done to determine whether a different profile market

shape exists for every commodity in every retail store or whether a single generalized profile shape can be developed.

4. More research needs to be conducted to determine the extent to which market area profiles vary with time.

5. It remains to be determined whether the least effort behavioral strategy which rational spatial action dictates is important in the choice of a particular center in a highly urbanized area. This question becomes particularly relevant given the duplication of retail facilities, the high overlap of market areas, and the high level of accessibility by consumers to retail facilities in highly urbanized areas.

6. A theoretical construct needs to be developed which will reflect the problem of variations in individual effort relative to the search for given types of commodities.

7. Research needs to be conducted to determine the extent to which an individual's travel behavior is in equilibrium with the objective structure of the city, and the effects of such variables as differences in race, social status, education, etc. on action space formation.

8. An effort is needed to develop, in theoretical terms, why the distribution of customers around a center should assume the shape of a normal curve, or perhaps more fundamentally whether such an occurrence actually happens.

9. We need to determine the effects of the consumer's subjective interpretation of the objective landscape on resulting patronage decisions in developing a more viable theory of consumer demand within the context of the spatial and temporal dimensions of market behavior.

#### FOOTNOTES

1. Joseph Barry Mason is Professor of Marketing and Urban Studies, Graduate School of Business, The University of Alabama.
2. Opportunities must be precisely defined in any use of the theory. The particular definition that is appropriate will depend on the type of social situation investigated, for distribution of opportunities for secretaries would necessarily be different from the distribution of opportunities for unskilled labor.
3. Carrothers has gone so far as to suggest that the distance exponent itself may be a variable which is inversely related to the size of the population and the distance itself (Carrothers, 1956). He also suggested that populations of different sizes should be raised to variable powers to reflect the effects of agglomeration economies.
4. Among other problems related to disaggregation and stratification is the problem of attaching exponents as weights to determine the attraction power between and among areas. In fact, this must be experimentally determined and, even though the computer may speed the process, the researcher is still confronted with the "fit" question.
5. Some studies contend that the yields of the gravity model should be viewed as theory and not as empirical regularity (Neidercorn and Bechdolt, 1969). Other studies have also established that Reilly's law on occasion yields results not in harmony with reality (Jung, 1959).

## REFERENCES

- Applebaum, W. & Cohen, S. The dynamics of store trading areas and market equilibrium. Annals of Association of American Geographers, 1961, 51, 81-83.
- Baskin, C. W. A critique and translation of Walter Christaller's Die Zentralen Orte in Suddeutschland. Unpublished doctoral dissertation, University of Virginia, 1957.
- Beckmann, M. On the equilibrium distribution of population in space. Bulletin of Mathematical Biophysics, 1957, 19, 81-90.
- Berry, B. & Horton, F. Geographic perspectives on Urban Systems. Englewood Cliffs, N.J.: Prentice-Hall, 1970.
- Berry, B. & Garrison, W. A note on central place theory and the range of a good. Economic Geography, 1958, 34, 304-311.
- Berry, B. Geography of market centers and retail distribution. Englewood Cliffs, N.J.: Prentice-Hall, 1967.
- Berry, B. & Pred, A. Central Place Studies: A bibliography of theory and applications.
- Bucklin, L. Shopping patterns in urban areas. Berkeley: University of California, Institute of Business and Economic Research, 1967.
- Bucklin, L. The concept of mass in intra-urban shopping. Journal of Marketing, 1967, 31 (4), 37-42.
- Bunge, W. W. Theoretical Geography. Lund Studies in Geography, Series C. General and Mathematical Geography, No. 1. The Royal University of Lund, Sweden: C.W.K. Gleerup, 1973, 136-137.
- Carey, H. C. Principles of Social Science. Philadelphia: J. B. Lippincott and Co., 1858.
- Carrothers, G.A.P. An historical review of the gravity and potential concepts of human interaction. American institute of planners, 1956, 22, 94.
- Christaller, W. Central places in southern Germany. Englewood Cliffs, N.J.: Prentice-Hall, 1966.
- Cox, R. Consumer convenience and the retail structure of cities. Journal of Marketing, 1959, 23.
- Fetter, F. A. The economic law of market areas. Quarterly Journal of Economics, 1924, 38, 520-529.
- Fite, G. & Reese, J. E. An economic history of the United States. Boston: Houghton Mifflin, 1965.
- Gambini, R., Huff, D. L., & Jenks, G. F. Geometric properties of market areas. Papers and Proceedings of Regional Science Association, 1968, 20, 85-92.
- Greenhut, M. The size and shape of the market area of the firm. Southern Economic Journal, 1952, 19, 37-50.

- Hagerstrand, T. The propogation of innovation waves. Lund Studies in Geography, Series B. Human Geography, No. 4. Sweden: The Royal University of Lund, 1952.
- Haggett, P. & Gunawardena, K. A. Determinants of population thresholds for settlement functions by the Reed-Muench method. Professional Geographer, 1963, 16, 6-9.
- Hoover, E. M. The location of economic activity. New York: McGraw-Hill Book Co., 1948.
- Horton, F. E. & Reynolds, D. R. Action space formation: a behavioral approach to predicting urban travel behavior. Travel Analysis. Highway Research Record 322. Washington, D. C.: Highway Research Board of the National Academy of Sciences, 1970.
- Hoetelling, H. A mathematical theory of migration. Unpublished masters thesis, University of Washington, 1921.
- Huff, D. L. A note on the limitations of intra-urban gravity models. Land Economics, 1962(a), 38, 64.
- Huff, D. L. A probabilistic analysis of shopping center trade areas. Land Economics, 1963, 39, 81-90.
- Huff, D. L. Defining and estimating a trading area. Journal of Marketing, 1964, 28, 36-38.
- Huff, D. L. A probabilistic analysis of consumer spatial behavior. In W.S. Decker (Ed.), Emerging Concepts in Marketing. Chicago: American Marketing Association, 1962(b).
- Huff, D. L. A topographical model of consumer space references. Papers and Proceedings of the Regional Science Association, 1960, 6, 159-173.
- Huff, D. L. Ecological characteristics of consumer behavior. Papers and Proceedings of Regional Science Association, 1961, 7, 19-28.
- Isard, W. Location and space economy: A general theory relating to industrial location, market areas, land use, trade and Urban Structure. Cambridge, Mass.: The M.I.T. Press, 1956.
- Isard, W. Methods of regional analysis. Cambridge, Mass.: The M.I.T. Press, 1960.
- Jung, A. F. Is Reilly's law of retail gravitation always true? Journal of Marketing, 1959, 24, 62-63.
- Losch, A. The economics of location. Translated by William H. Woglom and Wolfgang F. Stolper. New Haven: Yale University Press, 1954.
- Lynch, K. The image of the city. Cambridge: The M.I.T. Press, 1960.
- Mason, J. B. Theories of urban development: Comparative regional planning and development. Washington, D.C.: Government Affairs Institute and United States Department of State, 1973.

- Mason, J. B. & Moore, C. T. An empirical appraisal of behavioristic assumptions in trading area studies. Journal of Retailing, 1970-71, 46, 31-37.
- Mason, J. B. & Moore, C. T. The corridorized concept of household shopping highway travel behavior. Annals of Regional Science, 1970, 4, 69-80.
- Mason, J. B. & Moore, C. T. The reliability of census data in trading area analysis and store location studies. Mississippi Valley Journal of Business and Economics, 1969, 5, 68-72.
- Moore, C. T. & Mayer, M. L. The uses of gravity models as measures of retail trade area potential and the influence of interstate highways. Paper presented at the Southwestern Social Sciences Association Annual Meeting, Dallas, Texas, Spring 1966.
- Moore, C. T. & Mason, J. B. Major retail center patronage decisions. Journal of Marketing, 1969, 33 (3), 61-63.
- Mulvihill, D. & Mulvihill, R. Geography, Marketing, and Urban Growth. New York: Van Nostrand Reinhold, 1970.
- Neidercorn, J. H. & Bechdolt, B. W. An economic derivation of the gravity law of spatial interaction. Journal of Regional Science, 1969, 9, 273-282.
- Neyman, J. & Scott, E. L. On a mathematical theory of population conceived as a conglomeration of clusters. Cold Spring Harbor Symposia on Quantitative Biology, 1957, 22, 109-120.
- Nystuen, J. D. A theory and simulation of intra-urban travel. In W. L. Garrison and D. F. Marble (Eds.), Quantitative geography, Part I: Economic and cultural topics and dimensions of micro-behavior. Evanston, Ill.: Northwestern University, Department of Geography, 1967.
- Olson, G. Distance and human interaction: A review and bibliography. Regional Science Research Institute Bibliography Series No. 2. Philadelphia: Regional Science Research Institute, 1965.
- Ravenstein, E. G. The laws of migration. Journal of the Royal Statistical Society, 1885, 48, 167-235.
- Revzan, D. A geography of marketing: Integrative Statement. Berkeley: Institute for Business and Economic Research, 1968.
- Revzan, D. A geography of marketing: Resource bibliography. Berkeley: Institute for Business and Economic Research, 1968.
- Richards, P. Shock waves on the highway. Journal of the Operations Research Society of America, 1956, 6, 42-51.
- Schwartz, G. Development of Marketing theory. Especially Ch. 2 entitled "Laws of Retail Gravitation." Cincinnati, Ohio: Southwestern Publishing Co., 1963.
- Scott, P. Geography and Retailing. Chicago, Ill.: Aldine Publishing Co., 1970.
- Simons, P. L. The shape of suburban retail market areas; Implications from a literature review. Journal of Retailing, 1973-74, 49, 65-78.

- Stouffer, S. A. Intervening opportunities: A theory relating mobility and distance. American Sociological Review, 1940, 5, 845-857.
- Thompson, D. Future directions in retail area research. Economic Geography, 1966, 42, 6.
- Von Thunen, J. H. Der isoleite Staat: In Peter Hall (Ed.), Von Thunen's Isolated State. New York: Pergamon Press, 1966.
- Warntz, W. Potential of population and its relationship to marketing. In Reavis Cox and Wroe Alderson (Eds.), Theory in marketing. Chicago, Ill.: Richard D. Irwin, 1950.
- Weber, A. The theory of the location of industry. Translated by C. J. Friedrich. Chicago, Ill.: The University of Chicago Press, 1929.
- Zipf, G. K. The P,P<sub>2</sub>/D hypothesis: On the inter-city movement of persons. American Sociological Review, 1946, 11, 677-689.



A MULTIVARIATE ANALYSIS OF PREPURCHASE  
DELIBERATION AND EXTERNAL SEARCH BEHAVIOR

Stephen J. Miller and William G. Zikmund  
Oklahoma State University

Several studies have indicated that economic and demographic characteristics such as age, income, education and family life cycle, relate to household shopping behavior. Based on a major survey of household grocery shopping behavior, this study offers further support to such evidence. The study indicates that deliberation and search are not simple univariate phenomena but should be examined in a multivariate framework. Canonical analysis was utilized to test this multivariate framework. The analysis yielded profiles of households associated with a variety of deliberation and search activities.

Information evaluation and search activities in grocery shopping behavior is a topic of considerable interest to the public policy maker as well as the marketing manager. For example, consumer advocates recommend that disadvantaged consumers need to receive more and better product information. Marketing managers need to know the best information medium for each market segment. While many psychological factors, such as perceived risk and personality, are considered to be determinants of search behavior, research efforts show socioeconomic variables are also determinant factors.

Thorelli (1971) has observed, "There is no reason to assume that product information is either homogeneously or normally distributed among the population." Several research studies indicate that individuals are more responsive to information sources they perceive to be compatible with their own personal characteristics, such as family life cycle or social class (e.g., Engel, Kollat, and Blackwell, 1973; Katona and Mueller, 1955; Levy, 1969; Thorelli, 1971). One might expect upper middle class women, married less than ten years, to organize shopping more purposefully (less impulsively) and more efficiently than lower class women (e.g., see Kollat and Willet, 1967; Levy, 1969). Or increases in multi-store food shopping might be expected to correlate with multiple-automobile availability or education (Prasad, 1972).

While the evidence suggests that socioeconomic factors or personal characteristic factors do appear to influence the propensity to search, it has been pointed out that there is a lack of studies investigating a combination of variables (Engel, Kollat, and Blackwell, 1973, p. 383). For instance, the influence of income or education on search behavior might vary at different stages of the family life cycle. Furthermore, utilization of one source of information may be correlated with the use of other sources of information (Katona and Mueller, 1955). Thus, researchers investigating personal characteristics, and their relationship to shopping deliberation and information utilization, should consider the interaction effects within both sets of variables.



This research extends the inquiry into the nature of the relationship between prepurchase deliberation and search behavior and personal characteristics, by utilizing multivariate analysis. Prepurchase deliberation and search behavior might include a number of activities such as frequent shopping trips, the reading of newspaper advertising, shopping list preparations and multistore shopping. The view is taken in this study that the activities represent a total process of behavior rather than simply a number of discrete and independent variables. Likewise, socioeconomic factors or personal characteristics that identify segments differing by shopping process need not be examined separately or used to form a global index of social class, life cycle or the like. Investigating these variables, with their potential interactions as process and descriptor groupings in mind, should add to the understanding of the determinants of various search behavior.

The research reported in this study sought to answer the following question:

Are there profiles of personal characteristics of households that significantly relate to various prepurchase deliberation and search behavior processes?

The area of household grocery shopping was chosen as the focal point for investigation. This shopping activity has a variety of characteristics that made it amenable to the study. First, it is an activity in which all households participate. Second, the food budget represents a sizeable percent of the household expenditures which makes extensive shopping behavior an economically rewarding task. Finally, the food industry provides a large amount of marketing-dominated information.

The complete list of variables identified for the study is given in Table I. A variety of planning and search process variables reflect the criterion set. The predictor set of variables are those frequently examined individually as determinants of shopping behavior.

The data for the study was gathered by The Daily Oklahoman and the Oklahoma City Times newspapers in their 1972 consumer audit of the Oklahoma City SMSA. The sample included 2020 households utilizing an area-probability sample for the metropolitan area. Within each sample household, an in-home personal interview using professional interviewers was conducted with the "chief marketer," usually the housewife. Questions included brand purchase behavior, shopping behavior, media exposure characteristics, and household personal characteristics.

In the research, bivariate correlation is first used to examine simple relationships between the criterion and predictor variables. Then canonical analysis is employed to isolate profiles of personal characteristics and shopping behavior patterns. A canonical correlation maximally correlates two linear additive sets of variables measured across a group. For a more complete discussion of this statistical technique, see Alpert and Peterson (1972).

#### The Bivariate Results

A variety of studies have been cited earlier in the paper that dealt with an explanation of prepurchase deliberation and search behavior components

TABLE I  
Research Variables

Newspaper A Food Ad Readership	---Frequency of food ad readership in a large metropolitan morning newspaper (ADREAD A)
Newspaper B Food Ad Readership	---Frequency of food ad readership in a large metropolitan evening newspaper (ADREAD B)
Newspaper C Food Ad Readership	---Frequency of food ad readership in moderate size morning paper (ADREAD C)
Use of Shopping Lists	-----Frequency in use of preplanned shopping lists: 1 = use of list (LIST)
Store Variety	-----Number of food stores regularly shopped (VARIETY)
Trip Frequency	-----Number of separate trips to grocery store made by family (TRIPS)
Travel Distance	-----Distance one is willing to travel to favorite food store (DISTANCE)
Homeownership	-----Ownership of home: 1 = renter status (HOMEOWN)
Residency	-----Length of residency at current address (NEWCOMER)
Young Children	-----Number of children under age six (CHL6)
Older Children	-----Number of children age six to seventeen (CH617)
Adults	-----Number of adults in household (ADULTS)
Mans Age	-----Age of the male head of household (AGEMAN)
Womans Age	-----Age of the female head of household (AGEWOMAN)
Household Head	-----Sex of the head of the household: 1 = male (HEAD)
Marital Status	-----Marital status of head of household: 1 = married (MARRIED)
Family Education	-----Education level of the male head of household (EDUCMAN) education level of the female head of household (EDUCWOMAN)
Family Income	-----Total income for the family (INCOME)
Race	-----Race of head of household: 1 = white (RACE)
Car Ownership	-----Number of passenger cars owned by family (CARS)



(shopping behavior) based on personal characteristics. A bivariate correlation matrix for the shopping and personal characteristic variables is given in Table II. Many of the correlations between shopping and personal characteristic variables were statistically significant. However, the correlations were quite low with the highest being .20 for number of shopping trips and number of cars in family. Thus, only four percent of the variation in the shopping variable can be explained by transportation availability.

Another striking feature of the correlation matrix is the apparent low intercorrelation among the shopping variables. The highest correlation is .21 for a variety of stores shopped and number of trips. This moderate interaction indicates that a number of separate dimensions of the planning-search process of shopping are represented by the data.

A canonical analysis of the data should cast light on the various shopping patterns exhibited by households and the characteristics descriptive of such households. While bivariate correlations revealed few shopping-personal characteristic relationships, canonical analysis might indicate combinations of variables of the criterion and predictor sets that interrelate. Market segments based on shopping process variables will have been defined and identified.

#### The Canonical Analysis

The canonical analysis included the seven criterion variables and fourteen predictor variables as listed in Table I. Of the twenty-one variables, five "point" measures are included (list usage, home ownership, household head status, marital status, and race). The assumptions of canonical analysis include interval scaled data. However, as with similar multivariate techniques, the assumption can be relaxed with little damage to the analysis (Peterson and Alpert, 1972).

Seven canonical roots were generated by the analysis, one for each criterion variable. A summary of these are given in Table III along with relevant statistical tests of significances. For example, the first canonical

TABLE III

#### Canonical Correlation Roots

Canonical Root	Canonical Correlation	Chi Square	DF	Prob>x <sup>2</sup>
1	.41	770.28	98	.0001
2	.32	400.81	78	.0001
3	.22	192.50	60	.0001
4	.15	95.54	44	.0001
5	.12	50.16	30	.0120
6	.09	20.21	18	.3209
7	.05	5.54	8	.7002

R of .41 indicated that 16.8 percent of the variation in the linear combination of shopping variables could be explained by the linear combination of personal variables in the predictor set. Bartlett's "chi-square test" was

used to test the statistical significance of each canonical correlation coefficient in the table. In all, five of the seven coefficients were significant at the .05 level. Thus, it appears that shopping patterns do relate to personal characteristic variables.

The next step in the analysis was an examination of each statistically significant root to ascertain shopping patterns that emerged and the various personal characteristic profiles that relate to each pattern. As with factor analysis, the significant relationships are not always readily interpretable. However, potentially useful inferences can be drawn from the interpretation although they must be treated as exploratory. One could examine the standardized canonical coefficients for both the criterion and predictor variables to identify salient variables from each set. However, this approach to analysis can be misleading if intercorrelations exist within either set of variables. Such is the case for the predictor set where some moderately high bivariate correlations exist. An alternative method of interpretation is to examine the correlations between each variable of the criterion and predictor sets and its respective canonical variate (linear combination of all the variables in the set). The relative size of the loadings within a set indicate the general influence each variable exerts in defining the dominant group of variables displayed by the root being investigated. This form of interpretation is analogous to an examination of factor loadings in factor analysis.

The analysis of variate-variable correlations for each canonical relationship began with identification of the highest loadings (correlation). Variables were included in the interpretation of the equation in the order of their strength, stopping either when the loading was less than .25 or the new variable no longer "made sense" intuitively. Obviously, care is taken to avoid a posteriori rationalization to fit prior notions. The correlations of each variable with its respective criterion or predictor variate are provided in Table IV for the five statistically significant roots. Each root is explained separately below as patterns of shopping process and related profiles of personal characteristics are identified.

#### Canonical Root One

Although the criterion variables differ in the magnitude of their correlations, every shopping variable could conceivably be included in defining this form of shopping. Preplanning of shopping is conducted through grocery ad readership in a variety of papers and a shopping list is developed to reify the planning. Then, many trips are made to a variety of stores, some a distance away from the home, as the planned purchases are secured and comparison occurs. In summary, this shopping pattern epitomizes the "ideal shopper" in planning and search for market needs. An examination of the predictor set of variables provides a personal characteristic profile. The highly correlated variables are reflected in a marriage household with both the husband and wife having higher than average education, high income, and the husband being somewhat older. Multiple-car ownership emerges as provider of the mobility for comparative shopping. The personal characteristics description in this case reflects the general findings from previous studies. Better educated, higher income households (upper middle class) in moderate stages of the family life cycle purposefully organize shopping efforts.

TABLE IV

## Canonical Variates - Variable Correlations

Criterion Set (Shopping Behavior)	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>
Adread A	.47 <sup>a</sup>	-.36 <sup>a</sup>	.00	-.17	-.50 <sup>a</sup>
Adread B	.39 <sup>a</sup>	-.40 <sup>a</sup>	-.19	-.23	-.40 <sup>a</sup>
Adread C	.22	-.13	-.52 <sup>a</sup>	-.63 <sup>a</sup>	.51 <sup>a</sup>
List	.29 <sup>a</sup>	-.42 <sup>a</sup>	.67 <sup>a</sup>	-.14	.26
Variety	.60 <sup>a</sup>	-.36 <sup>a</sup>	-.11	.48 <sup>a</sup>	.37 <sup>a</sup>
Trips	.64 <sup>a</sup>	.42 <sup>a</sup>	-.29 <sup>a</sup>	.29 <sup>a</sup>	.00
Distance	.41 <sup>a</sup>	.44 <sup>a</sup>	.45 <sup>a</sup>	-.34 <sup>a</sup>	.04
Predictor Set (Personal Traits)					
Homeowner	.02	.43 <sup>a</sup>	.25	.11	.55 <sup>a</sup>
Newcomer	.24	-.52 <sup>a</sup>	-.43 <sup>a</sup>	-.05	-.06
ChL6	.06	.43 <sup>a</sup>	.45 <sup>a</sup>	-.30 <sup>a</sup>	-.15
Ch617	.42 <sup>a</sup>	.42 <sup>a</sup>	-.39 <sup>a</sup>	-.20	.04
Adults	.66 <sup>a</sup>	.15	.08	-.10	-.34 <sup>a</sup>
Ageman	.54 <sup>a</sup>	-.12	-.29 <sup>a</sup>	-.06	-.38 <sup>a</sup>
Agewoman	.06	-.71 <sup>a</sup>	-.20	-.34 <sup>a</sup>	-.03
Head	.64 <sup>a</sup>	.27 <sup>a</sup>	.14	.16	-.22
Married	.70 <sup>a</sup>	.24	.27 <sup>a</sup>	-.05	-.23
Educman	.59 <sup>a</sup>	-.01	.34 <sup>a</sup>	.28 <sup>a</sup>	-.21
Educwoman	.54 <sup>a</sup>	-.30 <sup>a</sup>	.41 <sup>a</sup>	.14	.28 <sup>a</sup>
Income	.74 <sup>a</sup>	.00	-.02	.18	-.01
Race	.33 <sup>a</sup>	-.10	.14	.71 <sup>a</sup>	-.20
Cars	.56 <sup>a</sup>	.23	-.09	.28 <sup>a</sup>	-.01

a: Variables that loaded on the variates at levels of .25 or higher and entered the variate interpretation.

Canonical Root Two

A quite different form of shopping appears in the second root. Here is seen the process involving virtually no preplanning through ad readership or list preparation. Instead, frequent trips occur, perhaps for some distance, to one or two stores. Little comparative shopping is done among stores and the lack of planning precipitates the need for frequent shopping. This pattern might be labeled as "unplanned shopper" in the market place. The household profile drawn from the loadings indicated a different family life cycle and educational level than in root one. Here is found the renter, new to the neighborhood, a young wife of lower education level, and children in the house of all ages. One sees the lower education and age combining to diminish the preplanning. This, combined with a larger family of children, necessitates frequent shopping.

### Canonical Root Three

The third shopping pattern to be considered introduces preplanning of one form but little comparative shopping. List preparation loads highly although food ad readership does not. Few trips are made to the shopping site a long distance from the home. In-house planning occurs while the shopping task is a minimum-effort process. The "one-stop shopper" or, perhaps the "in-store shopper" is illustrated in this case.

The personal characteristic profile that emerges for this root resembles the immediately previous root. Key variables reflect that the woman of the house is young, a renter new to the area, and has small children. However, in contrast to the previous root, there are no school-age children and the housewife's education appears higher. This profile reflects preplanning as suggested by the restrictive burden of young children hampering ability to comparatively shop and the housewife's better education.

### Canonical Root Four

Root four illustrates the type of shopping pattern where there specifically is no preplanning through ad readership in Paper C (inferring the newspaper's personality). Multiple trips are made to a variety of stores, yet a short distance traveled infers that any comparative shopping is in close proximity to the home. The personal characteristics indicative of this pattern are a white household with a young wife and no children. An automobile is available for mobility but does not enter the shopping process. A "neighborhood shopper" image appears as would be common in ethnic neighborhoods of many large cities. As an alternative interpretation, the behavior may be indicative of residents of a high retail density areas.

### Canonical Root Five

Here the variables reflect preplanned shopping through food ad readership limited strictly to Paper C and combined with list preparation although specifically not consulting the other newspapers. The only other variable that loads heavily is the variety of stores visited. Neither number of trips nor distance are salient characteristics.

The characteristics that correlate highly with this pattern indicate the single woman with no children, probably living alone, in rental housing. This cluster of individuals spans a number of age levels since it didn't enter the description. A label is somewhat difficult for this household although it appears to represent a "single-moderate planner" type.

## Conclusions

The findings from several studies unidimensionally investigating the role of economic and demographic characteristics on shopping behavior have indicated age, income, education, family life cycle, and other such variables as relating to shopping behavior. The results of the bivariate analysis in this study offers further support to such evidence. Our study indicated middle-aged female grocery shoppers are more prone to deliberate and to utilize shopping information. Individuals in earlier stages of the life cycle "shopped" less. Income was correlated with external search activity (distance traveled, number

and variety of trips). Education in this study, and in all the studies reviewed, appeared to be a consistent factor related to information utilization.

The study suggested that deliberation and search are not simple univariate phenomena. Further, the combination of variables generated for respective sets of criterion and predictor variables adds clarity to our understanding of their interacting influence. Certain households read food ads in many different papers while preparing a shopping list. Others read only specific papers or just prepared a list. In the search activity, some households shopped a variety of stores over a broad area. Others restricted shopping to a variety of stores in close proximity. The variety of personal characteristics that described the criterion variable appear to be related to prepurchase deliberation and search behavior. Variate one portrays a deliberate shopper, characterized by a mobile family higher in income, and education and in one of the later stages of family life cycle. Other combinations (e.g., variate 2 vs. variate 3) suggest that factors influencing the perceived cost of search (e.g., children and babysitter expense) are moderated by other demographic factors such as education.

This study offers evidence that the total search process should be viewed as a "system" of activities and that market segmentation, perhaps, may be more fruitful when two or more variables are utilized. Some groups of people emphasize in-store activities in their search process while other differing on socioeconomic factors use a much different process of gathering information. From a public policy point of view, it seems that better educated consumers as well as truthful product labeling information, is the most fruitful area for improving disadvantaged consumers' positions.

#### FOOTNOTES

1. The authors wish to thank the Oklahoma Publishing Company for providing the survey data for this study.
2. Stephen J. Miller is an Associate Professor of Marketing and William G. Zikmund is an Assistant Professor of Marketing at Oklahoma State University.

#### REFERENCES

- Alexis, M., Simon, L. S., and Smith, K. M. Some determinants of food buying behavior. In M. Alexis, R. S. Holloway and R. S. Hancock (Eds.), Empirical foundations of marketing. Chicago: Markham, 1969.
- Alpert, M. I., and Peterson, R. A. On the interpretation of canonical analysis. Journal of Marketing Research, 1972, 9, 187-192.
- Bucklin, L. Consumer search, role enactment and market efficiency. Journal of Business, 1969, 42, 416-438.
- Darden, W. R. and Reynolds, F. D. Shopping orientations and product usage rates. Journal of Marketing Research, 1971, 8, 505-508.
- Engel, J. F., Kollat, D. F., and Blackwell, R. D. Consumer behavior. New York: Holt, Rinehart and Winston, 1973, 156-157 and 373-392.



- Frank, R. E., Massy, W. F., and Wind, Y. Market segmentation. New York: Prentice-Hall, 1972.
- Katona, G., and Mueller, E. A study of purchasing decisions. In L. H. Clark (Ed.), Consumer behavior: The dynamics of the consumer reaction. New York: New York University Press, 1955, 30-87.
- Kollat, D. T. and Willett, R. P. Consumer impulse purchasing. Journal of Marketing Research, 1967, 4, 21-31.
- Levy, S. J. Social class and consumer behavior. In S. W. Newman (Eds.), On knowing the consumer. New York: John Wiley and Sons, 1968.
- Martineau, S. P. Social class and spending behavior. Journal of Marketing, 1958, 23, 121-130.
- Peterson, R. A., and Alpert, M. I. Consumer choice patterns and psychographics: A test of predictability. Proceedings of the Fourth Annual Meeting of the American Institute for Decision Sciences, 173-180.
- Prasad, V. K. Correlates of multistore food shopping. Journal of Retailing, 1972, 48, 74-81.
- Rich, S. U., and Jain, S. C. Social class and life cycle as predictors of shopping behavior. Journal of Marketing Research, 1968, 5, 41-49.
- Schapker, B. L. Behavior patterns of supermarket shoppers. Journal of Marketing, 1966, 30, 46-49.
- Sparks, D. L., and Tucker, W. T. A multivariate analysis of personality and product use. Journal of Marketing Research, 1971, 8, 67-70.
- Thorelli, H. B. Concentration of information power among consumers. Journal of Marketing Research, 1971, 8, 427-432.

## CREDIT RISKINESS OF LOW-INCOME CONSUMERS<sup>1</sup>

Donald E. Sexton, Jr.<sup>2</sup>  
Columbia University

Low-income consumers frequently experience difficulty in obtaining credit. In part, this situation is due to credit-granting procedures based on the general population. This investigation was an attempt to find variables that would identify those low-income consumers who are good credit risks. A sample of 4,119 credit histories from three national retailers was examined.

At present, the many low-income families who are unable to obtain credit must pay high prices to inner city merchants -- prices that include sizable charges for credit. If more reputable retailers can extend credit to the less affluent consumer, then perhaps the costs of consumer durables to the poor may be reduced. That was the hope with which this study was begun.

### The Data

Data were obtained through the generous cooperation of J.C. Penney, Montgomery-Ward, and Sears-Roebuck. However, to preserve confidentiality, those retailers are not identified elsewhere in this report.

The effective sample of 4,119 families was randomly selected from those families, of an original sample of 38,000, for which there was complete information on the variables used in the analysis. The sample was constructed so that it would contain a relatively large number of low-income families (under \$5,000 per year). All the data were gathered in 1972. The data from each of the three retailers came from a different city: Pittsburgh, Chicago, and Los Angeles.

Available for every one of the 4,119 families were the following variables.

- Evaluation of payment history (good or bad)
- Retailer where have account
- Married
- Single
- Divorced or separated
- Widowed
- Number of dependents
- Age
- Primary montly income (dollars)
- Presence of extra income
- Own home
- Rent home

Have telephone  
Credit investigation made

In addition, families for each of the three retailer pairs had other variables in common:

Sex  
Other income  
Second job  
Property income  
Occupation  
Months at present job  
Wife working  
Boarder  
Mobile home dweller  
Live with parents  
Months at present address  
Checking account  
Savings account  
Department store reference  
Loan reference  
Other reference  
Number of major and minor derogatories  
Previous account  
Retail or catalog customer  
Application date  
Highest monthly payment  
Current monthly payment

For retailer A, a bad credit risk was defined as a customer more than 150 days past due or one who was written off with a balance in excess of \$30. For retailer B, bad risks were those accounts in collection. For retailer C, those customers two months or more past due were labeled bad credit risks. These discrepancies in definitions were unavoidable. In the analyses, the retailer variable may be expected to remove the linear component of such variation. As a digression, there were many differences in variable definitions when the data were first received; it required eight months to make the three data sets compatible.

One major shortcoming to this study is that the data describe consumers who have already been granted credit. Therefore, the results may be biased toward a certain type of good credit risk. An ideal sample for this study would require a retailer to grant credit to everyone he could find. In the absence of such an experiment, one must use these data and extrapolate with caution.

The high-income family and low-income family samples were somewhat different in composition. High-income families were more likely to be married and to own homes. On the average, families in both groups had lived at their present addresses about the same length of time, but high-income workers, on the average, had been at their current jobs approximately six years longer than low-income workers. As one might expect, relatively more high-income families had loan references, checking accounts, and savings accounts. However, the percentages of those

families able to supply department store credit references and other types of credit references were the same for both high-income and low-income families.

### The Results

Because of the large sample size and the capacity of the computer programs available, it was necessary to employ regression analysis instead of discriminant analysis. The dependent variable was binary: 0 for a bad risk, 1 for a good risk. Separate regressions were made across all high-income families and across all low-income families (Table 1). For the high-income customers, most pairwise correlations among the independent variables were below .3. However, there were high correlations (over .7) between the married and single variables, the own home and rent variables, and phone and credit investigation made variables. In addition, the correlations indicated that store B customers were more likely to have phones and to not have had credit investigations. Among the low-income families, store B patronage, age, and monthly income were highly correlated, as were number of dependents, age, and store B patronage.

In Table 1 the regression coefficients are presented in the form of beta coefficients. These are standardized coefficients that can be directly compared both with other coefficients in the same regression and with coefficients in other regressions.<sup>3</sup>

TABLE 1

#### Credit Riskiness V. Consumer Characteristics

Characteristic	High-Income Families	Low-Income Families	Chow Test (F)
Retailer A	-.15	-.18	3.32
Retailer B	-.05	-.05	3.08
Married	.12**	-.01	7.38**
Single	.01	-.06	4.14*
Divorced or Separated	.01	-.07	5.22*
Number of Dependents	-.12**	-.10*	3.12
Age	.15**	.20**	3.08
Monthly Income	.06**	.03	3.24
Presence of Extra Income	-.03	.01	3.98*
Home Owner	.08*	.04	3.32
Home Renter	-.11**	-.07	3.74
Have Telephone	.22	.32	3.39
Credit Investi- gation Made	.23**	.31**	4.10*

N	3291	828	
R <sup>2</sup>	.58**	.57**	
F	341.1**	82.**	

\* p < .05

\*\* p < .01

Overall, the regressions in Table 1 provided good fits to the data. Moreover, the statistically significant coefficients generally had logical signs. Besides those shown in Table 1, regressions with various subsets of the independent variables were examined to avoid possible collinearity problems. However, those results were consistent with those in Table 1 and have been omitted for brevity.

Because the set of variables common to all three retailers consisted of only thirteen variables, additional regressions were made across samples composed of data from the various pairs of retailers. These analyses allowed examination of an additional twenty-five variables. All the regressions for the pairs of retailers exhibited close overall fits, but fewer than half of the regression coefficients were statistically significant and many of these were those variables found to be significant in the regressions across all three retailers. These results are available from the author, but have been omitted since they add little to the remaining discussion.

The main question of this investigation was: Do the variables associated with good credit risks differ between high-income and low-income families? Overall, the regression coefficients for the high-income families did not significantly differ from those for the low-income families.<sup>4</sup> However, an all-coefficient test is perhaps too broad and one might ask whether or not coefficients for individual variables differed between the high-income and low-income samples.

In fact, the individual coefficients were significantly different for few variables (Table 1). These variables consisted of: married, single, divorced or separated, credit investigation made, and presence of extra income. The married variable was strongly positively associated with good credit risks among the high-income families; the single and divorced or separated variables were strongly negatively associated with good credit risks among the low-income families. The credit investigation made variable was a useful forecaster of good credit risks for both groups, but was of more importance for low-income families. Finally, although the coefficient for presence of extra income differed between the high- and low-income samples, the beta coefficients were small, indicating it was not an important predictor variable of credit riskness.

### Summary And Conclusions

An analysis of a large sample of retail credit histories found variables that were useful predictors of credit riskness. However, the predictive abilities of only a few of these variables differed between high-income and low-income families. Based on this sample and the analysis to date, one cannot conclude that the identification of good credit risks among high-income and among low-income families requires different procedures. It is possible that further analyses examining possible interactions may yet find different predictor variables for high-income and low-income consumers.

## FOOTNOTES

1. This research was supported by the Graduate School of Business, Columbia University. The author wishes to thank Wallis Hocker, Ed Petti, and Ravinder Sharma of J.C. Penney, Morton Schwartz of Montgomery-Ward, Linden Wheeler and James Smith of Sears-Roebuck, and Robert Shay of Columbia University for their interest, cooperation and advice.
2. Donald E. Sexton, Jr. is Associate Professor, Graduate School of Business, Columbia University.
3. See Ferber for a detailed explanation.
4. For details of the Chow Test of coefficient differences, see Beckwith or Johnston.

## REFERENCES

- Andreasen, A. (Ed.) Improving inner city marketing. Chicago: American Marketing Association, 1972.
- Beckwith, N. E. Test of equality between coefficients in several linear regressions. Columbia University. Working paper, 1972.
- Durkin, T. (Ed.) Proceedings of the Arden House Conference on Consumer Credit. New York: Columbia University, 1972.
- Ferber, R. Market research. New York: McGraw-Hill, 1949.
- Johnston, J. Econometric models. New York: McGraw-Hill, 1963.
- Sturdivant, F. (Ed.) The ghetto marketplace. New York: Free Press, 1969.



SCARCITY AND HOARDING: ECONOMIC AND SOCIAL EXPLANATIONS  
AND MARKETING IMPLICATIONS

Ronald Stiff, Keith Johnson and Khairy Ahmed Tourk  
Illinois Institute of Technology

It seems unusual to experience hoarding in modern society, yet a number of products have been hoarded in 1974. Scarcity alone is insufficient to explain these hoards. Social and economic theories provide explanations for hoarding demonstrating its occurrence under specialized conditions not always requiring scarcity. Scarcity is but one of many signals available to consumers as an encouragement to hoard. In addition to retail availability hoarding is influenced by signals including price, formal and informal communications interacting with preconditions including consumer experiences and expectations. Theories of hoarding lead to marketing implications involving distributions, pricing, advertising and public relations decisions. It is concluded that if hoarding continues to exist in modern society the resultant secondary effects will both encourage and permit the development of empirically based theories of hoarding.

Consumers and managers in industrialized countries have come to expect that there will seldom be a scarcity of any goods during stable economic times. Although shortages may be expected for newly introduced products, shortages for most products are contrary to consumer and management expectation. Yet, in early 1974 shortages seemed to be epidemic. The news of shortages spread rapidly, aided by mass media sources which cited shortages in gasoline, toilet paper, toilet seats, toilets, paper bags, yellow tennis balls, anti-freeze and tomato paste: the latter attributed to an enormous increase in the taste for pizza in Japan (Hollie, 1974; Malcolm, 1974; Business Week, 1974). This atmosphere of scarcity has led to reports in the press of hoarding by consumers such as runs on toilet paper, "panic buying" of gasoline and in some cases stock-outs of products.

Hoarding in an affluent society has received only limited attention by behavioral scientists and has been of little interest to most marketers. Hoards in affluent societies have generally taken the form of collections such as art, oriental rugs, precious metals (including coins), currency (which may include stamps), and jewels. These hoards are of a different nature than current hoards of low priced consumer products. Collections are based on assumptions of long term scarcity of the collected goods, which are often of high intrinsic value and may replace other investments or currency during times of financial stress or war, while hoards of low priced consumer goods are based on expectations of short run instability of supply. Current hoards are affecting a larger number of consumers and producers than collections, which we will not consider.



## The Concept of Hoarding<sup>1</sup>

Hoarding exists when the consumer's current inventory of an item exceeds his inventory in previous periods while his expected consumption rate (taste) remains constant. This definition is operational for both consumers and organizations at both individual (micro) and group (macro) levels. The degree of hoarding can be represented as the ratio of current inventory to previous inventories. Hoarding also exists when the consumer's expected consumption rate is changing. This condition is more difficult to measure since changes in consumer inventory must be controlled for changes in expected consumption. Wartime hoarding often follows this pattern as the consumer seeks to both increase his inventory and decrease his consumption to make more effective use of stocks which he expects to decline.

The relationship between hoarding and the expected consumption is influenced by consumer expectations of the duration of shortages. If the consumer expects shortages to be short-run he may hoard without adjusting his consumption; however, expected long-run shortages are likely to lead to declines in expected consumption. These relationships suggest an outline for a theory of hoarding in which consumer preconditions, including expected consumption and expectation of shortages, lead to hoarding and hoarding produces secondary effects including changes in seller inventory (supply) and expected consumption. These secondary effects may amplify or dampen hoarding. For example, a visible decrease in available supply may amplify hoarding while a decrease in expected consumption may dampen hoarding.

Signals, such as visible supply, interact with consumer preconditions to produce hoarding. Economists consider price and total supply as signals influencing hoarding, while sociologists place greater stress on mass media and inter-personal communications and the consumer's direct observation of both retail availability and the purchase behavior of other consumers. A number of these signals are under the control of marketing managers, including advertising, price, and to some extent retail availability.

An outline for a theory of hoarding is shown in Figure 1. In this theory scarcity is only one of several possible signals leading to hoarding. The relationship between hoarding and scarcity is complex. First, hoards can exist without scarcity which apparently was the condition which existed for toilet paper in many parts of the United States in early 1974. Second, hoarding may lead to scarcity which allegedly was the condition for gasoline; the United States Federal Energy Office estimated that "rolling hoards" of gasoline were created by motorists who topped off their tanks, accounting for as much as a four day supply of gasoline nationally (Ingersoll, 1974). Finally, scarcity may lead to hoarding which is a condition that arises frequently during wartime and following natural disasters but may not have occurred yet in modern industrial society (Anatoli, 1971 and Salisbury, 1969).

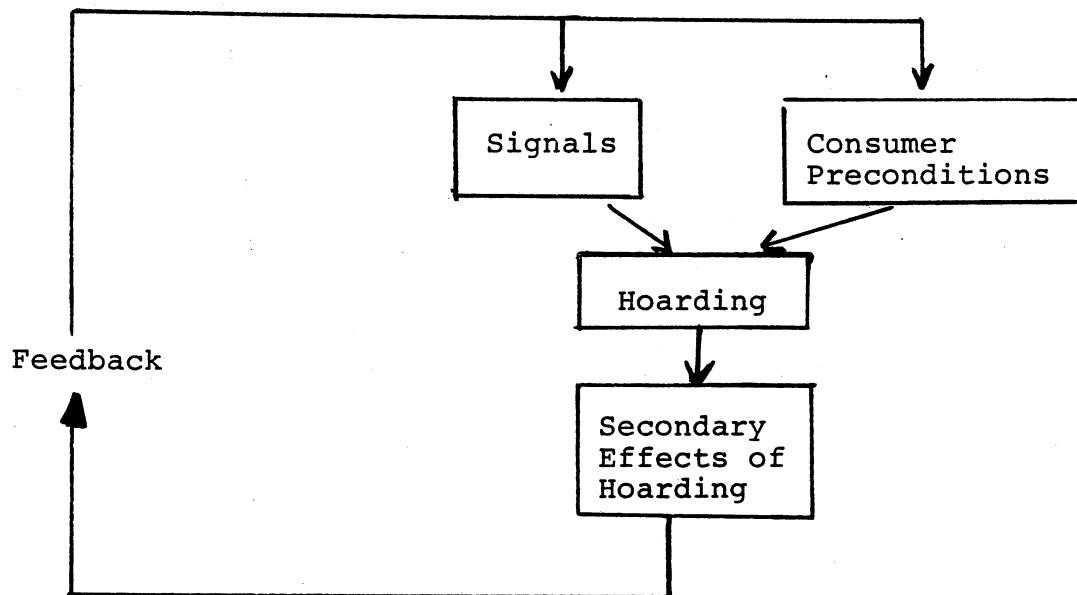


Figure 1: Outline of a theory of hoarding

#### Economic Explanation of Hoarding

Hoarding by consumers takes place in goods that have inelastic demand and low income elasticities. Total spending on these goods tends to be relatively small compared to total consumption expenditures. Hoarded goods generally have few substitutes and are relatively easy to store. For consumer hoarding to take place, two conditions should be satisfied. The first condition can be explained within a two-period framework. The motive to hoard in period  $t$  the goods that are normally purchased in period  $t+1$  depends on comparing two types of costs:

First: the cost of acquiring  $t+1$  goods in period  $t$  (i.e., "present" cost) and

Secondly: the cost of purchasing  $t+1$  goods in period  $t+1$  (i.e., "future" cost).

A necessary condition for hoarding is that the "present" cost is less than the "future" cost.<sup>2</sup>

The "present" cost does not include only the cost of buying the goods but also the cost of storing them and the cost of the tied-up funds. This type of cost could be called the objective cost. The reason for this is that the consumers generally know the interest foregone (if they use their own money to buy the goods), the present price of goods and their storage costs.

On the other hand, the "future" cost depends on the expected price of goods. Different consumers could have different expectations concern-

ing future prices. This makes the "future" cost a subjective cost. In calculating the future costs, the individual consumer is influenced by market and non-market signals. The consumer also takes into consideration the distinct probability that the hoarded good might eventually disappear from the traditional marketplace. Such an eventuality might give rise to the development of a black market. The extra cost of search and information associated with the black market should be considered as a part of the "future" cost.

The second condition for consumer's hoarding is that it is primarily motivated by a desire to secure his own supplies of the hoarded good rather than making financial gain. This last condition is what makes hoarding different than speculation. As with speculation it is true that hoarding provides a means of hedging against future price fluctuations and the individual consumer gains or loses depending on whether his expectations turned out to be correct or not. With hoarding, however, the change in the consumer's financial status is secondary to his primary objective of securing an adequate supply of the hoarded good for his own personal consumption.

Secondary economic effects result from hoarding. Hoarding is an increase in demand. In the short-run with supply fixed, the price of the hoarded good would increase. This price increase does not necessarily mean that a real scarcity problem exists in the long-run. In the long-run situation, the direction of price change is dependent on the supply conditions and the consumer's evaluation of these conditions.

The price of the hoarded good would decrease if the producers reacted to the increase in demand by increasing the supply of the good. The longer the time period under consideration and the higher the mobility of the inputs used to produce the good in question, the higher would be the elasticity of supply and the lower would be the increase in price in the long-run. The increase in the quantity supplied might also cause the consumers to revise their previous expectations concerning the hoarded good. A decision on their part to hoard less or to stop hoarding would reduce demand bringing prices down even further.

On the other hand, if the long-run elasticity of supply was low or zero, then the degree of hoarding might intensify causing demand to increase. With every increase in demand, the price of the hoarded good would continue rising producing inflation. For consumers living on relatively fixed incomes, inflation produces a reduction in their standard of living reducing their share of the income distribution. This effect is most severe when the hoarded good represents a significant part of the consumers income, especially among the poor.

The above conclusion is not necessarily universal. One could think of a case where goods were hoarded by the poor groups in the society. This is particularly true if poor people have more time at hand to invest in restructuring their purchasing behavior relative to hoarded goods. In this case if the rich are forced to pay a higher price for the same goods then, other things equal, a change in income distribution would take place in favor of the poor.<sup>3</sup> This, however, can be counteracted by the fact the rich are more willing to pay higher prices, have access to more information and more storage space to hoard the goods than the poor. If the shortage of goods continued, there would be

widening of the income gap between the rich and the poor in favor of the first. The outcome is a definite deterioration in the poor's standard of living. To deal with the problem a rationing system might be in order.

#### Social Explanation of Hoarding

Studies of social behavior show that people respond to information (signals), but signals themselves are not sufficient to induce behavior. Without necessary preconditions, even a strong signal will be ignored or reinterpreted, while a weak or ambiguous stimulus will be interpreted as a desired (or feared) signal with the necessary preconditions.

Studies showing the interaction of signals and preconditions include the 1938 War of the Worlds radio broadcast, in which music was interrupted by "news" bulletins stating that New Jersey had been invaded, "false" or spurious sounding of Civil Defense sirens in the 1950s, and a Texas fireworks factory explosion followed by a huge, dreaded mushroom cloud over the area. Preconditions at the time of these signals led to substantial collective behavior which would be unlikely today.

When a signal strongly indicates action and people have pre-conditioned expectations such as the possibility of life on Mars or enemy attack their response is to seek supporting evidence. (Baker and Chapman 1962). The War of the Worlds broadcast led to a large increase in telephone useage as did the false Civil Defense alerts. Frequent responses to these events included turning to others (co-workers, family, etc.), looking out the window for signs of anything unusual, turning on the radio and telephoning for information.

Often the search for new evidence leads to selective perception and interpretation. Some selectively interpret everyday conditions as frightening, as when traffic patterns lead some War of the Worlds listeners to believe in an "invasion" of New Jersey. Fewer cars than usual led one on-looker to suppose that people were in hiding or killed, while more cars led another to believe people were fleeing the "invasion."

The search for evidence can be effective in changing the information usually available to the individual. A signal may be simultaneous for many people, so when an abnormally large number of people pick up their telephone they find the lines gone dead, busy signals, or crossed lines - each leading to further fear and/or belief in the signal.

If small events have gone unnoticed, a signal may cause heightened awareness and receptivity for information previously ignored. An "epidemic" of auto windshield pitting has been traced to a newspaper article describing unusual pitting of a few windshields whereupon hundreds of motorists suddenly "discovered" they also had small defects in their windshields. Presumably, these defects had gone unnoticed for months or years in most cases.

In general, people are unlikely to engage in unusual behavior unless they have become sensitive to circumstances leading to that

behavior, and expect the conditions surrounding it. Panic (flight) does not occur in natural disasters, false air raid warnings, wartime, or other similar situations unless participants believe specific conditions are present requiring flight. These conditions consist of imminent danger, a belief that escape is possible, and a belief that chances of escape are steadily diminishing when present. These conditions lead to such events as runs on banks, stock market drops and theater panic with people pushing and shoving to escape (and endangering those unfortunate enough to fall or be crushed against an obstacle).

We can generalize from these studies of social behavior to the less understood area of hoarding. Hoarding appears to be a form of behavior which follows the pattern "panic" (avoiding a danger or threat), but the component of escape is replaced by a positive seeking of goods. A "buying panic" could partly resemble the run on the bank or "gold" rush.

For people expecting a normal, plentiful supply of goods, hoarding is unthinkable. Signals of an impending scarcity of goods alone do not induce hoarding. We suggest that the experience of being caught short is the chief pre-condition to consumer awareness of shortages. Further the objective costs of having run short, its dramatic qualities and its subjective interpretation by the consumer lead to generalized awareness of shortages which influence future buying decisions.

The direct costs of running short include:

- 1) expenses and time consumed in search for short goods in traditional and black markets;
- 2) added costs required in seeking a substitute; and
- 3) inconvenience caused by absence of short goods and inadequacy of its substitute.

The dramatic qualities of the experience of running short are a product of the salience of the good and the visibility of its shortages for the consumer.

The visibility of a product shortage is dramatically portrayed in businesses abandoned due to lack of supply, long waiting lines, social discussions with others about a shortage, and to varying extents, in the actual absence of goods.

Finally the consumer may assign the responsibility for a shortage to himself, to other consumers, or to powerful agents (government, manufacturers, or suppliers). If he accepts the blame (in not purchasing enough personal inventory), the consumer is more likely to hoard and be sensitive to future signals of on-coming shortages than if he blamed other consumers (who caused the shortage by hoarding).

Blaming official agents may increase a consumer's tendency to hoard, but it will also increase his suspicions of signals coming from such agencies.<sup>4</sup>

By all these factors American consumers received an experience likely to change their orientation towards goods from the shortages of gasoline last Winter.

The general unavailability of gasoline was accompanied by large increase in price, long waiting lines, (which paradoxically caused much gas to be wasted), closed stations, foregone vacation trips and unused vacation homes, and anxious moments concerning the family car, which involves the greatest household investment after the home.

The widespread reporting of shortages of various consumer products, frequently accompanied by a rush to buy in 1974 can be interpreted to be due to a large extent to consumer preconditioning by the 1973-74 gasoline shortage, combined with signals for individual products.

The bizarre social behavior in relation to consumption in 1974 (anger, hoarding, panic buying) must be understood as due to abnormal social conditions. Consumer expectations of abundance were shaken by experience with shortages. If products return to abundance, expectations will return to normal. Expectations may also normalize if the availability of products stabilizes at levels of intermittent scarcity, or even chronic scarcity, as in conditions of famine or war, although these expectations will be quite different from those of plentiful supply and may include hoarding as a matter of course.

It is the period of uncertainty (transition) from one condition to another which has extreme forms of reaction. One type of reaction (hoarding) may become normal in relation to goods in short supply while when it appeared in 1974 hoarding frequently seemed bizarre.

The role of signals has also undergone a period of uncertainty and misinterpretation for years. When consumers have become sensitive (preconditioned) to scarcity, they will read statements different from their interpretation during times of abundance. An official who pleads with consumers to stop panic buying will actually spread news of a shortage and cause increased buying. The message "There is no shortage of" will be interpreted "I didn't realize there is a developing shortage of" and "adequate supplies of - are on the way" will mean "there are problems in supplying."

In short the social explanation of hoarding is based upon the precondition of scarcity in the experience of the consumer. When scarcity is believed present, signals will be perceived (perhaps in error) that the supply of a good is uncertain and a rush to gain protection will tend to follow. Erratic hoarding behavior will accompany conditions of general uncertainty of the supply of good.

while predictable scarcity will cause more stable hoarding. Economic and social factors are shown in figure 2.

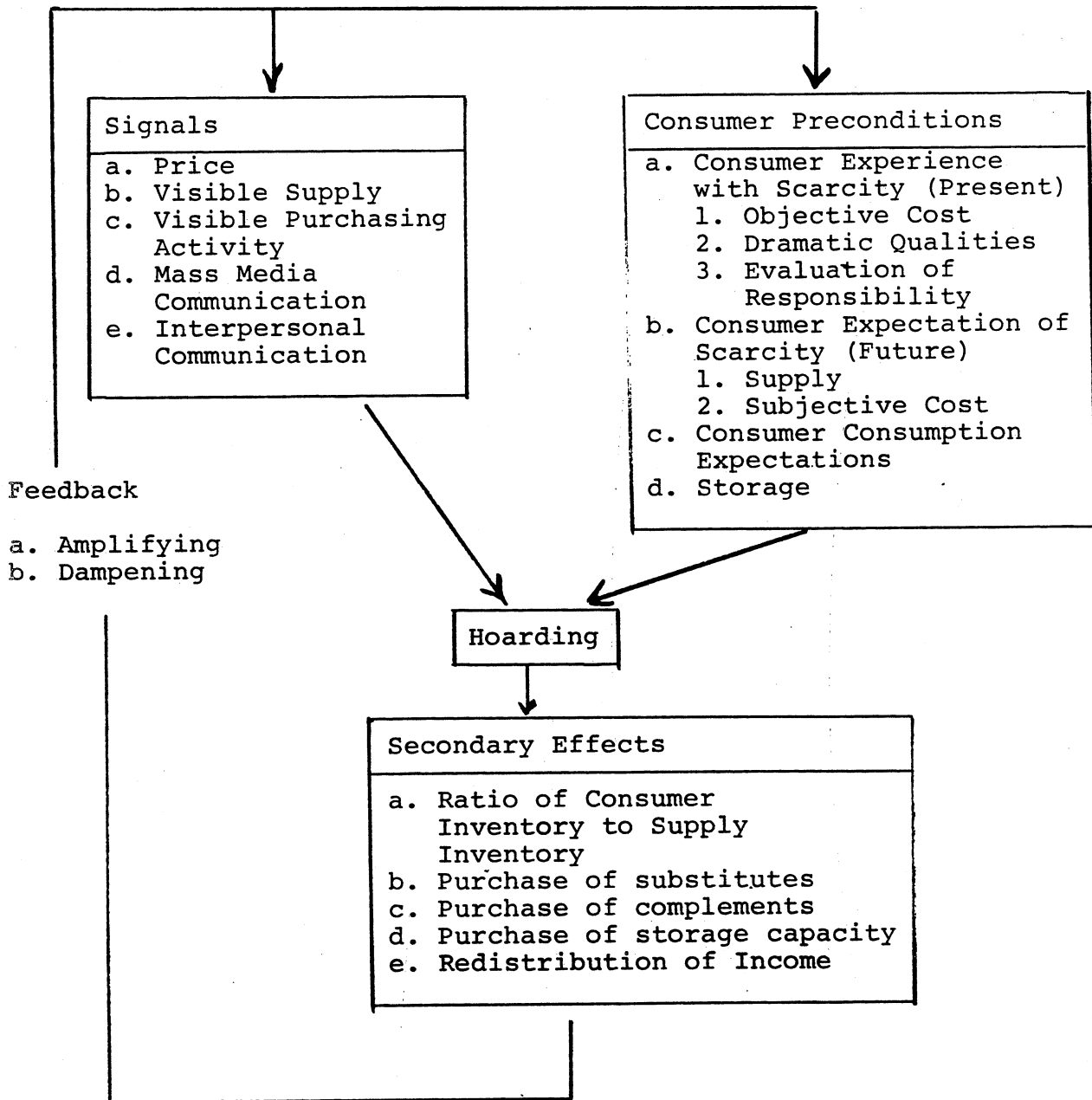


Figure 2: A theory of hoarding.

## Marketing Implications

These explanations suggest both threats and opportunities for marketing managers. The results of hoarding tend to create relatively uncontrollable market dynamics requiring both defensive and offensive marketing strategies.

### Market Dynamics

Once a product leaves a manufacturer it may be inventoried at the wholesale, retail, or consumer level. Hoarding acts to increase the consumer proportion of this inventory. For many products consumers have greater collective inventory space and purchasing power for a product than retailers and wholesalers combined. Thus for these products a relatively modest increase in consumer purchase rates may lead to little or no retail availability of a product or brand within a market area. As the consumer observes decreases in retail availability for a product or brand he will under appropriate preconditions view this as an early signal of shortages and increase his purchasing rate increasing the problem of retail availability. As individual brands of products stock-out the consumer is forced to either switch brands, switch retail locations, purchase a substitute product whenever possible, or to decrease his consumption of the product.

Ehrenberg (1972) has demonstrated that repeat buying for many non-durable products including breakfast cereals, coffee, gasoline, and toilet paper follow predictable statistical distribution. He proposes that this model is effective in predicting repeat buying except when there are "real differences in product-formulation, or price, or retail availability." Hence brands subject to stock-outs would not follow his theory of repeat buying and would be expected to lose market share position since brand loyalty would be threatened. Hoarding, therefore, may result in changes in market shares.

Ehrenberg, however, also proposes that when brand loyalty is very strong the consumer will revise his retail shop loyalty and seek other sources for the brand. Thus the retailer is likely to lose a part of his retail market share as consumers learn new shopping strategies. This may have happened for the purchase of gasoline in 1974.

An alternative to brand switching or retail shop switching is to seek out a substitute product. Thus those unable to purchase gasoline may switch to public transportation, those needing tomato paste may develop a taste for frozen pizza, cloth napkins may be substituted for paper napkins, and shortage of toilet paper may increase the demand for mail-order catalogues. As a result of the consumer learning to use substitutes, the demand for a product is likely to decrease, perhaps approaching zero. Thus we have the strange condition of a rapidly increasing demand (the hoard) leading to a shortage with resultant decreased demand for the product as demand for the substitute product increases.

If the consumer believes the supply of a product is seriously threatened he is likely to take steps to reduce his consumption of the product. This behavior as with product substitution leads to a long-run decrease in demand for the product. Hence, some persons who decreased their pleasure driving in current months may never return to this activity at their previous level. New modes of recreation, such as do-it-yourself activities may replace the threatened activity.



## Marketing Strategies

Marketing strategies which can be used to discourage hoarding or diminish the effects of hoarding on the organization's profits involve distribution, price, and advertising decisions.

Distribution strategies. It is likely that the first knowledge a marketer will have of hoarding by the consumer is the rapid decrease of retail inventory in some geographic areas. Since decrease in retail availability of brands and products are a signal increasing the likelihood of hoarding the marketer may be able to avoid stock-outs if he can move more of his product into the hoarding areas. If marketers cannot do this consumer feared shortages may well become a self-fulfilling prophecy as increasing consumer inventories exhaust wholesaler and retailer inventories. Whenever the marketer stocks-out in an area it is likely that he will either lose market share or force some of his retailers to lose retail market share.

This latter problem suggests a strategy for retailers. Retailers should be aware of products and brands that are likely to lead to changes in retail loyalty in the event of a stock-out. Inventory policy should place emphasis on control of these items. Shoppers can be polled as they check out to determine shortages, threats to retail loyalty, and hoarding. "Rain checks" can be given to encourage a wait and return attitude. At the same time retailers should place greater emphasis on services which encourage consumer loyalty. Thus gasoline service stations maintaining services such as washing windows and checking oil during gasoline shortages are likely to develop strong long run retail loyalty.

The success of distribution strategies are based on the ability to predict hoarding in time to reallocate supplies of the product. The short-run costs of reallocating supplies in the distribution system to cope with hoarding must be balanced against the potential long-run costs of stock-outs and for more fundamental financial reasons there is a great desirability of reducing distribution lead times.

Price strategies. We cannot think of any justification for decreasing prices in the face of hoarding; however, it is not clear when prices should be increased. It is generally assumed that increases in price will decrease demand. Once hoarding has begun an increase in price may be interpreted by consumers as a signal of future shortages and of future higher prices. Thus price increases may increase demand. A company is likely to increase its short-run profits by raising prices during hoarding, but a more realistic strategy is to use price as an instrument for maintaining retail inventory in order to minimize brand switching. The optimum strategy for a single company may be to meet competitor's price increases with slightly higher increases; however, this may trigger increased hoarding and retail shortages for many companies. Actual consumer behavior under these conditions of supply and pricing has not been reported.

Advertising and public relations strategies. Advertising Age (1974) has suggested "that in the future, when a rumor touches off doubt and panic, those advertisers whose products are involved must rush in with advertising - individually or collectively - to carefully advise the public of the true situation. Fast authoritative action on the part of advertisers and news media can head off trouble. To let rumor run rampant and do nothing to counter it is irresponsible." This strategy is likely to be counterproductive as a signal to consumers when consumer preconditions for hoarding exist. Counter-advertising may be viewed at worst as a signal of shortages to come or at best create additional interest in the product and heightened interest in observing retail availability. We are not convinced that any advertising will act to decrease demand except perhaps, that communicating negative characteristics of the product itself. We have too little experience with hoarding to know the influence of counter-advertising on the demand function, but it seems reasonable that positive product advertising should be decreased or eliminated when hoarding preconditions exist to improve organizational performance.

Only a company with extensive product inventory could benefit by extensive advertising. Advertising would increase generic product demand, leading to competitor's stock-outs, produce brand switching, and be likely to increase the short and long-run market share of those companies with available product inventory.

Advertising has been proposed as a perverse strategy as the following examples suggest:

"A widget maker decides that widget sales need a boost. He issues a statement to the effect that there is no widget shortage at present and none is foreseen in the future. It figures that in the absence of additional information, hoarding of widgets can begin and sales boom." (Advertising Age, 1974)

"At least one Chicago store is advertising that there is no shortage of record players, apparently hoping that this will prompt a run on record players." (Malcolm, 1974).

These signals will be ineffective when preconditions for hoarding do not exist. Additionally perverse strategies are more likely to increase short-run profits than long-run consumer confidence in the companies using them.

One strategy seems clear, marketers of substitute products should increase their marketing activities in the face of hoards. Advertising of substitutes should be increased and price changes should be carefully appraised. For example, the Chicago Transit Authority reduced prices by about one-half on Sundays for four weeks and advertised this action heavily during the 1974 "gasoline crisis". Ridership increased substantially although the long-run financial impact is unknown.

The general long-run strategy of a marketer is to reduce the number of signals which lead to hoarding. The most significant of these signals is retail availability. When availability is threatened strategies for price changes are not clear and should be viewed as highly situational. Advertising and price increases for hoarded products may lead to short-run increases in profits; however, longer-run profits are likely to result from a great reduction in advertising and careful consideration of price.

### Conclusions

This paper arises from concern over the current crisis psychology and its mix of preconditions and signals leading to hoarding. If shortages and hoards are a temporary problem then this paper is of temporary value. On the other hand, if future shortage and hoarding occur this paper serves to point out how little we understand consumer behavior underlying hoarding phenomena and the actions marketing management should take. If future hoarding is expected steps should be taken to improve our early warning systems or predictive methodologies. An opinion research corporation has advertised their public opinion poll as an "early warning system" presenting a picture of cars lined up at gas pumps. In addition to polls we could also use measurements of product flow at retail and wholesale levels and content analysis of news reporting as components of early warning methodologies. If hoards become more prevalent prediction methodologies will become more valuable and, of course, easier to test empirically.

Finally we have provided qualitative economic and social explanations of hoarding and subjective marketing strategies. A more quantitative theory of consumer hoarding behavior with more specific strategy decision rules would not be difficult to develop if more hoarding is experienced, thus increasing the demand for such a theory and permitting its testing.

### FOOTNOTES

1 This paper emphasizes the hoarding of products by consumers. Extensions of hoarding concepts and theory to both organizations and the purchase of services could be useful. Useful results might follow if services are considered as hoards in time or spatial dimensions rather than in inventory dimensions, (for example, being the first to see a new movie, going to Yellowstone park while it exists in its present state, or "excessive" preventive medical examinations).

2 A proper rate of discount is implicitly assumed here. This rate would allow direct comparison between "present" and "future" costs.

3 For example the poor may be more willing to spend greater travel time on public transportation and avoid the effects of hoarding and shortages of automotive products.

4 Blaming other consumers would seem to lead to consumer support for rationing, blaming self to hoarding, and blaming powerful officials to legal controls - even nationalization of producers.

## REFERENCES

- Advertising Age. Panic mongering and marketing. February 25, 1974, 16.
- Anatoli, A. (Kuznetsov) Babi Yar. New York: Farrar, Strauss and Giroux, 1971.
- Baker, George W. and Chapman, Dwight W. (Eds.), Man and society in disaster. New York: Basic Books, 1962.
- Business Week. Now tomato paste feels the squeeze. February 16, 1974, 50-51.
- Ehrenberg, A.S.C. Repeat buying: Theory and applications. New York: American Elsevier, 1972.
- Hollie, Pamela G. Here's a timely lesson on supply, demand, paper bags and things. Wall Street Journal. February 3, 1974, 1.
- Ingersoll, Bruce. FEO aid raps drivers for topping off. Chicago Sun-Times. March 1, 1974, 30.
- Malcolm, Andrew H. The shortage of bathroom tissue: A classic study in rumor. New York Times. February 3, 1974, 29.
- Salisbury, Harrison E. The 900 days: The siege of Leningrad. New York: Harper and Row, 1969.
- Smelser, Neil J. Theory of collective behavior. New York: The Free Press, 1964.



AN ANALYSIS OF MOVIEGOERS BY LIFE STYLE SEGMENTS<sup>1</sup>Glen Homan<sup>2</sup>, Robert Cecil<sup>3</sup>

Leo Burnett U.S.A.

and William Wells<sup>4</sup>

Needham, Harper, and Steers

The data in this study were used to demonstrate two points: the advantages of using life style data in addition to traditional demographics; and the advantages of sub-segmentation within a heavy user group. In the first phase of this analysis, heavy moviegoers were compared with nonmoviegoers on the basis of demographics and life style. In the second phase, the heavy moviegoers were separated into three groups via cluster analysis. These three groups were also compared on the basis of demographics and life style.

## Introduction

The findings reported in this paper are presented as empirical illustrations of two methodological points: (1) the advantages of using life style or "psychographic" information in addition to traditional demographic measures and (2) the advantages of sub-segmentation within a heavy user group. The paper shows that life style studies give a much fuller descriptive picture of the consumer than do demographics alone, that life style data may yield important dimensions that demographics alone may miss, and that life style data can give some idea of how products fit in the context of consumers' lives. It further shows that sub-segmentation both avoids the assumption that all heavy users are alike, and tells which variables really "go together" to characterize distinct consumer groups.

The data to illustrate these points come from a national study of 3700 married males, conducted by the Leo Burnett Co. in 1973. Respondents were members of the Market Facts Consumer Mail Panel. The sample is not intended to represent the entire movie-going population; it excludes married females and all single people.

The analysis was performed in two stages. In the first stage "heavy moviegoers" -- defined as those who had gone to the movies at least nine times in the past year -- were compared with nonmoviegoers - or people who had not gone to the movies at all in the last year. The groups were compared both in terms of demographics and in terms of life style, so as to give an idea of the differences between the two kinds of information. The life style analysis was based on two types of items: a series of 150 attitude, interest, and opinion statements (AIO's) on which the respondent indicated his degree of agreement or disagreement on a six-point scale; and a set of 56 activities on which the respondent indicated the number of times he participated in that activity in the past year.

The second stage of the analysis was a sub-segmentation by means of cluster analysis, among heavy moviegoers. This analysis produced three

segments of heavy moviegoers, described below. The purpose of this analysis was to show the advantages of sub-segmentation within a heavy user group.

Comparison of Heavy Moviegoers (N=758) versus Nonmoviegoers (N=1,122)

Demographics. Heavy moviegoers differ from nonmoviegoers on a number of demographic dimensions:

Heavy moviegoers are younger:

	<u>Non</u>	<u>Heavy</u>
Under 25	1%	15%
25-34	7	32

Heavy moviegoers are better educated:

	<u>Non</u>	<u>Heavy</u>
College graduate	8%	19%
Post-grad. degree	5	8

Heavy moviegoers have higher incomes:

	<u>Non</u>	<u>Heavy</u>
\$15,000-\$24,999	14%	27%
\$25,000 or more	3	7

Heavy moviegoers are more likely to live in an urban or suburban setting; nonmoviegoers are more likely to live in rural areas:

	<u>Non</u>	<u>Heavy</u>
Central City*	14%	22%
Suburban*	17	30
Rural	31	14

\*In SMSA of 500,000 or greater

Heavy moviegoers are more likely to be in professional or managerial occupations; nonmoviegoers are more likely to be craftsmen, farmers, or retired:

	<u>Non</u>	<u>Heavy</u>
Professional/Manager	26%	36%
Craftsman, etc.	25	20
Farmer	10	3
Retired	25	3

Heavy moviegoers tend to belong to larger households.

	<u>Non</u>	<u>Heavy</u>
2 member household	53%	27%
3 member household	18	24
4 member household	14	25
5 or more member household	15	23

Life Style. As with demographics, a number of life style differences arise between heavy- and non-moviegoers. Numbers in charts are percent agreeing with each statement (i.e., answering 4, 5, or 6 on 6-point scale).

Heavy moviegoers are more ambitious and optimistic. They have a richer fantasy life and are more self-confident:

	<u>Non</u>	<u>Heavy</u>
My greatest achievements are still ahead of me.	49%	80%
Five years from now the family income will probably be a lot higher than it is now.	63	88
I expect to be a top executive within the next 10 years.	11	37
I would like to take a trip around the world.	44	76
I'd like to spend a year in London or Paris.	19	42
I would do better than average in a fist fight.	33	55
I like to be considered a leader.	60	82
We will probably move at least once in the next five years.	24	49

Heavy moviegoers are more active and more socially-oriented:

	<u>Non</u>	<u>Heavy</u>
I like parties where there is lots of music and talk.	37%	65%
I like to play poker.	30	56
I would like to take a lesson in my favorite outdoor sport.	39	66
I always have the car radio on when I drive.	49	82

Heavy moviegoers have more "swinging" interests:

	<u>Non</u>	<u>Heavy</u>
"Playboy" is one of my favorite magazines.	13%	48%
I think I'm a bit of a swinger.	13	33
There are situations where sex outside the marriage can be a healthy thing.	15	35

Heavy moviegoers are more inclined toward new and sporty possessions:

	<u>Non</u>	<u>Heavy</u>
The new styles turn me on.	22%	46%
I like sports cars.	29	58

Nonmoviegoers have more traditional values:

	<u>Non</u>	<u>Heavy</u>
The U.S. would be better off if there were no hippies.	73%	46%
Liquor is a curse on American life.	59	36
I often wish for the good old days.	63	43
A woman should not smoke in public.	62	36
There is too much violence on TV today.	77	57
If Americans were more religious, this would be a better country.	83	63

Nonmoviegoers are less secure and more pessimistic:

	<u>Non</u>	<u>Heavy</u>
I don't like to take a chance.	71%	51%
I don't like to fly.	40	25
I am really looking forward to retirement.	75	58
A person should stay home whenever there is the slightest symptom of illness	55	28



Activities. Heavy moviegoers participate more often in virtually all activities monitored - except attending church:

	<u>Avg. No. of times participated in past 12 months</u>	
	<u>Non</u>	<u>Heavy</u>
Went bowling	3.4	9.8
Had a cocktail before dinner	11.4	18.5
Cooked outdoors	10.6	19.3
Listened to a tape or record	22.5	42.5
Attended a sporting event	4.3	10.3
Served wine with dinner	3.1	8.0
Went swimming	3.7	11.4
Brought home work	7.4	14.5
Attended church	33.2	25.3

Conclusion. Life style information enables us to obtain a richer and more complete portrait of the heavy moviegoer than the information provided by demographic data alone. While some of the life style differences could have been inferred from the demographic profile, it seems highly unlikely that all of them could have been.

#### Analysis of Three Heavy Moviegoer Segments

The second stage of the analysis was to employ a cluster analysis program that groups people together who respond in similar ways to a set of items.<sup>5</sup> In this case the items used were the 31 life style items on which the heavy moviegoers showed the most variance. It seemed likely that if heavy moviegoers do indeed consist of several subtypes, these types would be revealed by analyzing the items upon which heavy moviegoers agreed least.

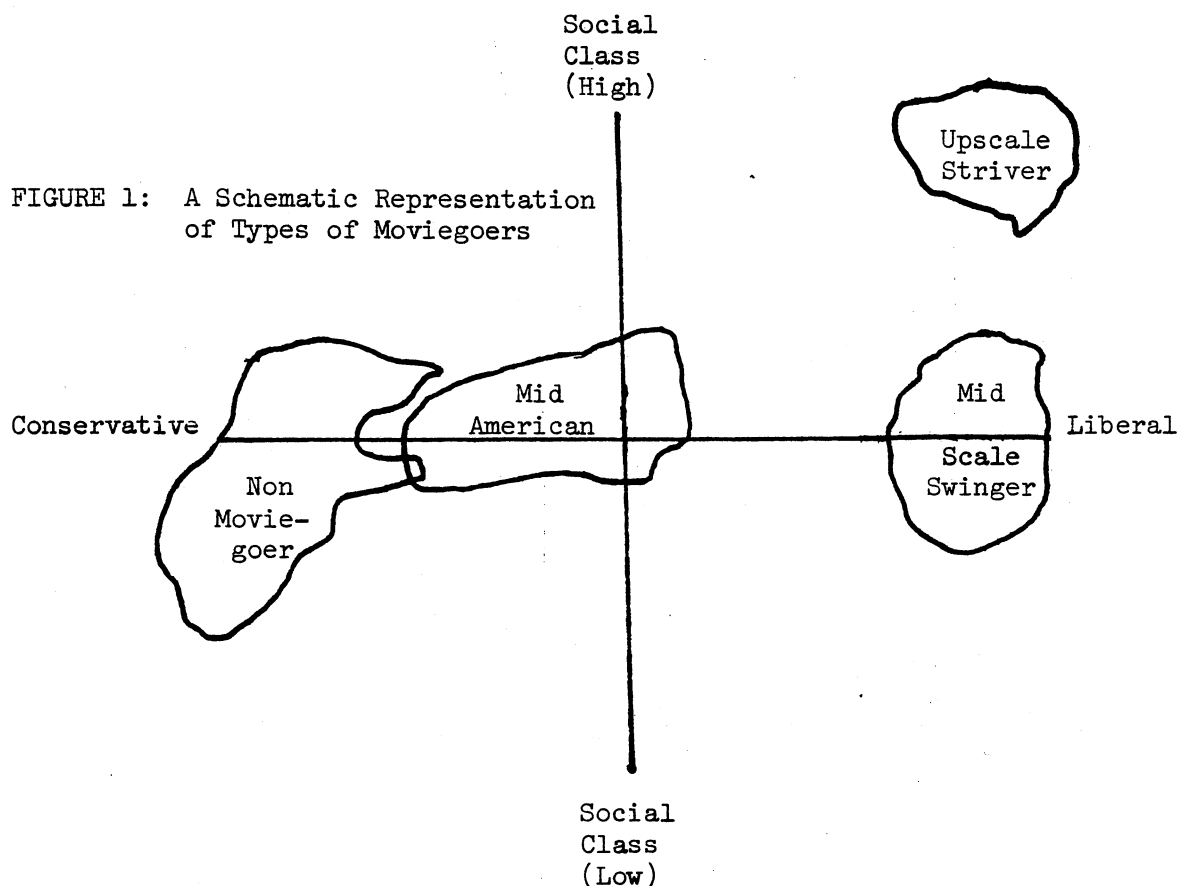
Membership in a given cluster is determined by the respondent's own answers, and is not predetermined by the researcher. The number of groups the program finds is predetermined by the researcher, however; so the researcher must compare several groupings to determine which is best. The criteria for selecting a solution were that each group should be coherent, internally consistent, reasonably large, and stable in a split-half test. A further criterion was that the cluster solution chosen should explain variance in the data better than differences produced by other cluster solutions. The three group solution fit these criteria. The groups were named the Mid American, the Mid Scale Swinger, and the Upscale Striver. They were examined for demographic differences and for differences on life style variables. Our sample included 272 Mid Americans, 216 Mid Scale Swingers, and 270 Upscale Strivers.

The demographic comparisons elicited some clear differences among the subtypes of heavy moviegoers. Level of education completed by the Upscale Strivers was much higher than that for any other group. The pattern of family income was similar: the Upscale Strivers were highest.

The nonmoviegoers are clearly older than any of the three heavy moviegoer groups; sixty one percent of them are over 54 years of age. The Upscale Strivers are somewhat younger (their mean age is 39); and the Mid Scale Swingers are the youngest group (70% are under 35; 25% under age 25). The Mid Americans, however, show no clear trends by age, being evenly spread across

all age categories. The Mid Scale Swinger is most likely to live in an apartment (although only 16% of them do). Nonmoviegoers are least likely to live in an apartment (only 3% do). Nonmoviegoers and Mid Americans are much more likely to live on a farm than the other groups; Upscale Strivers are more likely to own more than one car.

Life Style. When looking at both life style and demographic differences, the differences between groups are most simply explained by the following model:



The vertical axis in this model is social class, which was reflected in the demographics, and is also reflected in the answers to a number of the life style questions. The horizontal axis represents a number of attitudinal dimensions not covered in the demographic description; the labels "liberal or conservative" summarize a variety of dimensions including attitudes toward sex, attitudes toward churchgoing, and political orientation. The model shown above is an oversimplification because it doesn't show the changes in relationships between groups on some of the attitudinal dimensions. These changes in relationships will become clearer as we look at each of the groups in more detail.

First we will look at areas that differentiate nonmoviegoers from all types of heavy moviegoers. The earlier analysis assumed that if there was an overall difference between heavy moviegoers and nonmoviegoers, the difference would hold true for all types of moviegoers. This analysis is designed to show that there may be more than one type of heavy user of a product; and

that different types of heavy users may behave differently (even like non-users in some cases).

One area where nonmoviegoers were different from all three types of heavy moviegoers was in their view of the role of women. All three types of heavy moviegoers were more willing to accept changes in women's roles than were nonmoviegoers; as the following items show:\*

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
A woman's place is in the home	72%	59%	53%	49%
A woman should not smoke in public	62	42	32	34
Women wear too much make-up these days	64	53	52	46

\*The percentages shown for all the AIO's in this paper represent agreement with the items, i.e., 4, 5, or 6 on a six-point scale between "Definitely Disagree" and "Definitely Agree."

We also found that nonmoviegoers are more traditional in a number of other ways, on items ranging from family finance to the role of the church. The following list illustrates a few of these areas:

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
To buy anything other than a house or car on credit is unwise	62%	52%	46%	44%
There are day people and night people; I am a day person	85	76	49	62
Liquor is a curse on American life	59	47	30	27
I often read the Bible	46	34	17	22
There is too much violence on TV today	77	61	51	60

We found that all three types of moviegoers engage in many more activities than nonmoviegoers, as is illustrated by the following items:\*

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
Played board games	4.9	7.1	9.1	10.1
Went swimming	3.7	8.6	11.4	15.0
Finished a book	9.5	15.4	12.5	19.2
Washed the car	10.7	17.1	19.3	15.8

\*The scores on activities are derived from a seven point scale measuring the number of times each respondent engaged in an activity in the past year. The figures shown here are the mean number of times members of each group engaged in a given activity in that year. The categories had different intervals, however; these "means" are approximations.

This tendency extends across a wide range of items; the list could be extended if space permitted presentation of more data.

The final area of differences between nonmoviegoers and all three types of heavy moviegoers was in degree of self-confidence. All three types of heavy moviegoers are more self confident than nonmoviegoers, as these items show:

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
I think I have a lot of personal ability	82%	92%	96%	97%
Five years from now the family income will probably be a lot higher	63	79	92	92
A person should stay home whenever there is the slightest symptom of an illness	55	32	28	20
I often wish for the good old days	63	46	49	33

The heavy moviegoers have more generalized faith in themselves.

In summary, nonmoviegoers are different from all three types of heavy moviegoers in several respects. They are older than any of the groups of heavy moviegoers. They are also more traditional, less willing to accept changes in the role of women, less active, and less self confident.

The Mid American. The Mid Americans lie between the nonmoviegoers and the other two heavy moviegoer groups on the life style items. This is not to say that the Mid Americans do not have a distinct identity; the range from nonmoviegoers to the Mid Scale Swingers and/or Upscale Strivers is so great on many items that there is considerable room to fall in between. The Mid Americans use this latitude: In some areas their scores are similar to the scores of the other types of moviegoers; in other areas they resemble the nonmoviegoers.

Both the nonmoviegoers and the Mid Americans are more conservative about sex than the Mid Scale Swinger or the Upscale Striver:

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
There are situations where sex outside the marriage can be a healthy thing	15%	12%	51%	48%
"Playboy" is one of my favorite magazines	13	21	72	52
There is too much emphasis on sex today	86	83	48	57
You can't have any respect for a girl who gets pregnant before marriage	41	29	13	13
I like to think I'm a bit of a swinger	13	14	45	41

This finding is particularly interesting when viewed with one of our earlier findings. The Mid American is more like the other heavy moviegoers in accepting changes in the role of women, but at the same time disagrees with them sharply on being liberal about sex. The Mid American also bounces back and forth between these two poles on a number of other items relating to traditional values and personal conservatism.

We have shown some of the items where the Mid American is similar to the other heavy moviegoer types; here are some items where the Mid American is

more similar to the nonmoviegoers:

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
If Americans were more religious, this would be a better country	83%	74%	57%	60%
Young people have too many priveleges	79	72	54	61
A party wouldn't be a party without liquor	25	30	46	44
The new tyles turn me on	22	30	59	54
The U.S. would be better off if there were no hippies	73	60	35	42

We also find that the Mid Americans have more limited aspirations than the other heavy moviegoers:

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
I would like to have my boss's job	17%	20%	54%	60%
My greatest achievements are still ahead of me	49	67	89	88
Every man should have a den or a place for privacy	76	64	89	84

Further, the Mid American doesn't engage in much fantasy, and is generally rather cautious:

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
I would like to own the most expensive things	34%	38%	56%	57%
I would like to own and fly my own airplane	21	28	41	53
I would like to be an actor	6	5	22	22
I like danger	14	18	35	31
I don't like to take chances	71	67	56	36

Finally, the Mid American does in fact lead a very stable, settled life, with a definite routine:

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
We will probably move at least once in the next five years	24%	29%	54%	67%
In the last ten years we have lived in at least three different cities	11	10	24	35
My days seem to follow a definite routine such as eating meals at a regular time, etc.	76	73	54	53
I am really looking forward to retirement	75	67	55	50

In summary, when we compare the Mid American to the Upscale Striver and the Mid Scale Swinger, we find that the Mid American is more conservative about sex, has more traditional values in a number of other areas as well, and has more limited aspirations. The Mid American is also more physically and mentally settled into a routine. He actively rejects both fantasy and major changes in his life (except retirement), and doesn't like taking risks.

The Mid Scale Swinger. The items that define the Mid American and non-moviegoers at one end of a scale also define the Mid Scale Swinger and the Upscale Striver - at the other end. Stated briefly, the Mid Scale Swinger and Upscale Striver share the following characteristics: they are more physically mobile, lead active fantasy lives, don't follow a set routine, have high aspirations, and have more liberal attitudes toward sex, the role of women, and most other traditional American values.

The question that now arises is: what items differentiate the Mid Scale Swinger from the Upscale Striver, and which areas characterize each particular group? We already know a few demographic differences: The Upscale Striver has quite a bit more education and income, while the Mid Scale Swinger is younger. A number of life style variables differentiate between the two groups as well. The first of these is that the Mid Scale Swinger is very low on social and community activities, as shown in the following chart:

Activities	<u>Non Movie-goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
Did volunteer work	3.9	6.2	2.9	5.5
Attended Church	33.2	30.4	16.7	29.8
Collected paper, glass, etc. for re-cycling	3.6	3.9	2.2	5.2
Gave or attended a dinner party	5.4	6.1	4.8	13.8
AIO's				
I am or have been active in the PTA	27%	28%	18%	30%
Our family is a close-knit group	88	88	63	86

As a matter of fact, the Mid Scale Swinger is even lower on these activities than is the Mid American.

The Mid Scale Swinger is also materialistic, and somewhat dissatisfied and mistrustful.

	<u>Non Movie-goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
Most people are honest	78%	72%	64%	76%
No matter how fast our income goes up we never seem to get ahead	54	57	65	53
Spiritual values are more important than material things	79	72	60	70
I hate to get up in the morning	42	39	56	47

Finally, the Mid Scale Swinger is more likely to do manual repairs on his house and car than the Upscale Striver:

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
I do a lot of repair work on my car	46%	50%	62%	38%
We always use professional plumbers, carpenters or electricians for repair work around the house	39	33	21	41
Washed the car	10.7	17.1	19.3	15.8
Did a clean-up project around the house	15.1	14.5	12.8	17.9

He is, however, less likely to help with clean up projects around the house.

The Upscale Striver. Now let us turn to the areas that differentiate the Upscale Striver. First, the Upscale Striver is very hard working. He views his job as fulfilling and as a challenge, and he is very involved in his work.

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
My job requires a lot of selling ability	25%	28%	26%	66%
I expect to be a top executive within the next ten years	11	15	35	67
In my job I have to travel a great deal	21	25	23	60
In my job I do essentially the same things every day	73	72	72	44
Brought work home	7.4	7.2	5.2	35.2

The Upscale Striver is also the most active group in community affairs.

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
I would pay \$15 more on my income tax for a program of Federal pollution control	39%	49%	47%	60%
I've helped collect money for the Red Cross, United Fund, March of Dimes, etc.	34	34	30	44
I am interested in politics	56	50	55	77
Gave a dinner party	5.4	6.6	4.8	13.8
Returned an unsatisfactory product	2.1	2.4	2.7	4.1

The Upscale Striver is also consistently highest on activities for personal enrichment and entertainment (except TV):

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
Went to school	2.4	4.7	5.7	13.7
Jogged	1.4	2.3	3.9	8.1
Went out to dinner at a restaurant	16.4	20.7	19.9	35.9
Went to the library	3.0	4.9	3.5	11.5
I enjoy talking with my friends about the stock market and other investments	27%	30%	26%	56%
Television is our primary source of entertainment	69	59	63	48

This group, in short, is tremendously active in many different areas. This finding is consistent with the finding that the Upscale Striver has very high self esteem and considers himself a leader in his community:

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
My opinion on things doesn't seem to count very much in today's world	71%	63%	58%	45%
I would probably be a good politician	20	19	31	46
My friends and neighbors often come to me for advice	48	48	52	64
In my job, I tell people what to do	51	54	63	85
I am influential in my neighborhood	43	37	36	50

In addition, the Upscale Striver is innovative, and enjoys challenge and excitement in his life:

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
I often try new brands before my friends and neighbors do	41%	48%	55%	66%
I would rather live in or near a small town than in or near a big city	90	82	75	60
I often wish for the good old days	63	46	49	33
It might be better if we all returned to a simpler style of life	85	78	78	65

In summary, the Upscale Striver is hardworking, very involved in his job, and yet also very involved in his community and in his own personal enrichment and entertainment. The Upscale Striver is high in self esteem, and views himself as a leader; he is also innovative and likes challenges and excitement.

#### Conclusion

One of the purposes of this analysis was to show some of the types of information that life style items can add to the traditional demographics. The portraits of the groups were intended to illustrate the depth of information found in using life style items as opposed to using a demographic analysis alone.

The difference between life style analysis and demographic analysis should be particularly clear in the distinction between the Mid American and Mid Scale Swinger, where the only demographic difference is a slight difference in age. When life style items are used to compare the two groups, however, very clear differences emerge. Some of these differences, like the more traditional attitudes and orientation toward sex of the Mid American, might be posited from the slight difference in age. But we also find that the Mid American has more limited aspirations, engages in very little fantasy, and prefers (and has established) a life with much more routine. These



differences cannot be expected on the basis of a slight difference in age. The Mid Scale Swinger, on the other hand is much lower on community and social activities, and is more materialistic, mistrustful, and dissatisfied with life. These differences cannot be projected from the demographic profile alone.

These data also illustrate that treating heavy users of a product or a service as a single group may greatly oversimplify the data. It is possible to conclude that there is a difference between heavy users and nonusers when that difference does not exist for one or more of the subgroups within the heavy user group. The following items illustrate the point:

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
My job requires a lot of selling ability	25%	28%	26%	66%
In my job I do essentially the same things everyday	73	72	72	44
Most people are honest	78	72	64	76
I try not to eat foods that are high in cholesterol	56	51	38	56
Went on an out-of-town business trip	3.6	3.7	1.4	19.5

On the other hand, it is also possible to conclude that there is no difference between heavy users and nonusers - when in fact the heavy user group resembles nonusers only because it consists of several subgroups pulling in opposite directions. That phenomenon is illustrated by the following examples:

	<u>Non Movie- goer</u>	<u>Mid American</u>	<u>Mid Scale Swinger</u>	<u>Upscale Striver</u>
I do a lot of repair work on my car	46%	50%	63%	38%
Classical music is more interesting than popular music	37	38	26	42
I am good at fixing mechanical things	70	75	75	57
I always use the seat belt, even for a short drive	35	33	28	46
Collected paper, glass, etc. for recycling	3.6	3.9	2.2	5.2
Did volunteer work	3.9	6.2	2.9	5.5
Went hunting	3.4	6.2	5.5	2.4

The only way to get around these problems is to try one or more sub-segmentations. It is our belief that these sub-segmentations are more likely to be successful if one uses a statistical technique which groups respondents on the basis of their own patterns of response, rather than on the researcher's a priori ideas. That, however, is the topic of another paper.

## FOOTNOTES

1. Data, staff, and computer time for this research were supplied by Leo Burnett U.S.A.
2. Glen Homan is Account Executive, Leo Burnett U.S.A.
3. Robert Cecil is Assistant Research Analyst, Leo Burnett U.S.A.
4. William Wells is Director of Corporate Research, Needham, Harper and Steers
5. The specific technique used in the Leo Burnett cluster analysis program is a sift-and-shift technique developed at Market Facts, Inc. by Rich Johnson and described in a speech given February 15, 1972 for the New York Chapters of the American Statistical Association and the American Marketing Association, and entitled "How Can You Tell If Things Are Really Clustered?"



BRAND CHOICE AND COMMUNICATION BEHAVIOR:  
CONSUMER RESPONSE TO THE MARKETING OF RELIGION<sup>1</sup>

Donald J. Hempel<sup>2</sup> and William J. McEwen<sup>3</sup>  
University of Connecticut

In applying marketing concepts to the examination of a selected social institution, this paper reports results of a study of consumers of formal religions. Product perceptions and a range of consumer activities and attributes were surveyed in a mail sample of New England households. Segmentation of consumers by "product" and "brand" related values and activities was undertaken to provide feedback to promotion strategists. Emphasis is placed on personal values, evaluation criteria, and product-related communication behaviors of three identifiable consumer segments. Implications for broadening the scope and application of marketing theory are discussed for both the consumer behavior theorist and the social institution marketer.

Recent concerns for broadening the concept of marketing have resulted in a growth of application of marketing concepts to such nonbusiness areas as population control, health services and fund raising (Kotler, 1972; Kotler and Levy, 1969). Contrary to the conclusions presented by Tucker (1974), this new direction seems to hold considerable promise both for theory development and an improved responsiveness of social institutions. Basic marketing concepts (e.g. market segmentation) can be refined and improved through empirical testing in a broader range of social settings with a broader range of social "products." Theory which is based upon rather traditional conceptualizations of market transactions can perhaps be made more relevant to contemporary concerns (e.g. the corporate social audit) by integrating marketing concepts into a broader perspective of exchange relationships within large scale social systems.

The goal orientations and emphases on performance evaluation among nonbusiness organizations have certainly been influenced by management training seminars using business applications concepts. Systematic evaluation of marketing applications has, however, been lacking in many of these social institutions. It thus seems desirable to further evaluate Kotler's concept of generic marketing by examining the applicability of marketing concepts to organization-client relationships in those nonbusiness areas which Tucker suspects have little comparability (religious, political and educational organizations).

This paper examines the organization-client relationship for one of these institutions, viewing church-member interactions as a marketing problem. The study described was designed to provide regional leaders of a national church organization with information about consumer (nonmember as well as member) activities, interests and opinions as a means of both assessing the present status of and targeting future directions for their communication programs. The aspect of consumer behavior which was of particular interest here is brand choice -- more specifically, what are the

more useful bases for segmenting the religious markets of major concern to different denominations? The need for information of this type, as well as its potential use as a basis for directing religious communication strategies, has been recently summarized by Engel (1973).

The audience for religious communications appears both large and highly segmented, as evidenced by the number and variety of religious denominations represented in the array of local churches. Two market segments from this audience appeared to offer particular promise as client groups of special interest to those planning persuasion programs: (1) the church member (or "active consumer") group; and (2) those households which are in a transition state in terms of community affiliations (movers who are establishing themselves in new neighborhoods). Such segments can of course be additionally refined by introducing segmentation bases employed in more traditional product marketing settings. More notable (and heuristic to both theorist and strategist) would be extent of product usage (frequency and duration of church attendance), product involvement (including communication activities related to the product), and the various life style dimensions (e.g., Plummer, 1974).

#### Methodology

The study was conducted among samples drawn from three populations of New England households during the period of May to July of 1974. First, a sample of 949 church members was selected from the membership rolls of 30 churches in the New England Synod, Lutheran Church in America. The sample represented a multistage probability sample of all members listed in the current directories of the 140 synod churches. A second sample of 200 designated church leaders was obtained from a synod headquarters list of leaders in the 140 member churches. Finally, a sample of 960 recent home buyers was selected from the localities chosen for the church member sample (including communities adjacent to the selected towns). This probability sample was selected to represent a key market segment consisting of movers who purchased a home in the residential areas served by the 30 parochial units included in the church member sample.<sup>4</sup>

Each selected household received a cover letter, introducing the project as a study of how people in New England feel about churches and religious activities. The letter made no mention of denominational affiliation and stated explicitly that respondents need not be active church members. The study was presented as part of a continuing program of communications research undertaken by a team of university faculty members. The instructions stated that the questionnaire could be completed by any adult member of the household.

The four-page study questionnaire was divided into three parts.<sup>5</sup> Part I employed Likert-type scales to measure attitudes toward the role of churches, with emphasis on personal value systems and the meaning of religion. Part II measured opinions regarding the type of church-related information desired by respondents, their use of media for general local news and for church-related information, and the choice criteria they would advocate using for selecting a local church. Part III obtained basic demographic and life-style data, along with information about church attendance, religious affiliation and involvement in church activities.

Most questions were structured so that responses could be recorded with a single check mark. The selection of topics and wording of questions was based on previous studies of religious attitudes (Strommen et al., 1972) and on three focused group interview sessions conducted by the authors. The preliminary sessions indicated that many people (particularly those uninvolved in religious institutions), are hesitant to engage in discussions of their personal religious beliefs and values. This hesitancy and the length of the questionnaire were reflected in response rates lower than comparable studies of consumer attitudes toward commercial products. Usable questionnaires were returned by just over 27% of those who had been sampled, with response rates differing somewhat for the three populations represented. Refusal to answer specific value or demographic questions averaged about 4%.

### Results and Discussion

Extending a marketing approach to religious decision-making suggests at least two decision-outcomes of interest: product (religion) acceptance and brand (denomination) selection. Assessing the relationship of product and brand usage to information-seeking patterns and to attitudinal, attribute preference and life style patterns suggests a further breakdown in the presentation of results. Results are thus presented according to two defined subareas of interest: (1) the relationship of personal values and attitudes to consumer attributes and information preference patterns; and, (2) the relationship of product and brand involvement to personal values and to preferred religious ("product") attributes.

The results presented here involve a classification of the respondents on the basis of responses concerning their religion, mobility, and involvement in church activities. Those who did not respond to all three of these questions were eliminated for this analysis, even though the population from which they were selected was identifiable through color-coded questionnaires. This adjustment provides a set of more precisely defined market segments for comparison. The residual group of respondents (with missing data on any of these three classification variables) will be treated later as a separate market segment.

#### Personal Values and Consumer Attributes

Table 1 presents the intercorrelations of nine personal value variables with six consumer attribute variables for each of the three market segments. Coefficients are shown only for those variables where the correlation had a chance probability below .10 for at least one of the segments. Due to missing data on individual questionnaire items, the correlations are based on pairs of 95 respondents in the leader segment, 175 in the established church member segment, and 198 in the mover segment.

The first three personal value variables are index measures composed of several questionnaire items which had high loadings on the same orthogonal factor. Each of the remaining variables represent responses to single questions which the factor analyses indicated were relatively independent of the other measures or had high loadings on several factors (e.g., the personal importance of religion). All of the personal value measures used Likert-type scales.

TABLE 1

Correlation of Consumer Attitude/Values  
Attributes And Activities

Values/Attitudes	Attributes and Activities					
	Age	Educ. Level	Moved How Long Ago?	Moved How Far?	Church Attend. Freq.	Extent Of Church Involve.
1. Religious Dependence Index	.41	-.32	.23		.35	.09
	.45	-.23	.14		.36	.16
	.26	-.30	.01		.57	.14
2. Personal Irrelevance Index		-.04		-.29	-.19	-.32
		-.18		-.11	-.27	-.25
		-.01		-.08	-.29	-.01
3. Life Satisfaction Index					.15	
					-.05	
					-.09	
4. We really don't need all this concern with change	.22	-.26	.16	-.13		
	.15	.07	.17	-.05		
	.14	.10	-.11	.19		
5. Churches should speak out on issues of the day		.09			.01	.16
		.06			.21	.22
		.16			.08	.09
6. I get much more out of sermon than from any formal liturgy	.08	-.09	.03	-.11		
	.17	-.14	.21	-.15		
	.02	-.17	.03	-.06		
7. My religion is very important to me	-.06	.10			.31	.04
	.17	.09			.45	.30
	.10	-.17			.63	.21
8. Most religions are pretty much alike		-.17	.26	-.24	-.18	
		-.07	.22	.07	-.10	
		-.05	.00	.05	-.13	
9. Should seek information about different religions when moving	.17	-.16	.22	-.31	.00	
	.04	-.08	.09	.09	-.15	
	-.02	-.04	.01	-.02	-.24	

Note: The three items in each cell represent the Pearsonian bivariate correlation coefficients between the attitude/value measure and the attribute/activity measure for church leaders, established church members (non-leaders), and recent movers respectively. Data are presented only for those cells in which the correlation was significant at the .10 level for at least one of the three market segments.

Several relationships appear to merit attention as bases for the segmentation presumed to exist. As might be expected, religious dependence (traditional view of religion; religion as security for self and family) is positively related to age and church attendance, and negatively correlated with education. Age was also correlated with conservatism (no need for concern with change), as was stability of residence. The personal irrelevance of religion (money emphasis, insufficient social activities, too little to say about life) was negatively correlated with two measures of product usage--frequency of attendance and (for church members only) the degree of church involvement. The frequency of church attendance was highly correlated with the personal importance of religion for all market segments.

There are several other relationships which are interesting because they were unexpected, and/or because they provide additional bases for segmentation. Agreement with the statement that religions are pretty much alike was positively correlated with the stability of residential environment among the church members. This suggests that the market segment which might be regarded as the foundation of a church's position in the community (i.e., members listed in directory) may be less likely to perceive their religion as differentiated from others. The concern generated by this finding is partially offset by the finding that the frequency of church attendance is negatively correlated with the attitude that one should try to find out more about different religions when moving into a new community. In other words, frequent product usage appears to generate denominational (brand) loyalty.

Table 2 presents the intercorrelations between the personal value variables and the three measures of sources preferred for information about churches and church activities. Each index is composed of several measures of source helpfulness which factor analysis indicated to be closely related along the same dimension. Examination of differences in value structure according to media preferences reveals that religious dependence was greater among the member and leader segments who relied upon church sources (church newspapers, magazines, bulletins, and pastors). Recent movers who relied upon these sources were somewhat more likely to express satisfaction with their progress toward fulfilling personal goals. Reliance upon public media (local radio, newspapers, and television) and informal personal sources (relatives, friends, and neighbors) was positively correlated with religious dependence among church leaders and members. Recent movers who found these sources to be helpful were less likely to exhibit religious dependence either in terms of the general index of dependence or the specific rating of religion importance. The personal relevance of religion was lower among those who relied upon informal personal sources in all three market segments. Ratings of this source were positively correlated with life satisfaction among the leader and mover segments, but not among established church members.

In general, the ratings of information source helpfulness were not highly correlated with personal value structure. The findings do suggest that leaders and members who have greater religious dependence are more likely to rate all information sources as helpful in finding out about churches and their services. This relationship appears to be the opposite for movers--high dependence was associated with lower source ratings. One explanation for this pattern is that religious orientation fosters information search which increases consumer sensitivity to information available through all communications media. Recent movers with a strong, traditional religious orientation may be frustrated in their efforts to find information (particularly from public media) because they have less knowledge of where to search in local media.



TABLE 2

Correlation Between Consumer Attitude/Values  
and Sources of Information Used for Church Activities

Attitude/Value	Sources of Information About Churches and Church Activities		
	Church Source Index	Public Media Index	Informal Interpersonal Index
1. Religious Dependence Index	.16	.15	.19
	.23	.26	.29
	-.06	-.20	-.12
2. Personal Irrelevance Index			.21
			.20
			.21
3. Life Satisfaction Index	.04		.21
	.10		.07
	.17		.26
4. Religion very important to me		.08	.15
		.10	.04
		.18	-.20
5. Churches should speak out			
6. Get more out of sermon than liturgy			-.10
			.05
			-.18
7. No need for all this concern with change		.02	
		.15	
		-.07	
8. Most religions are alike			
9. Consider different religion when moving			.17
			.02
			-.10

Note: The three items in each cell represent the Personian bivariate correlation coefficients between the attitude/value measure and the attribute/activity measure for church leaders, established church members (non-leaders), and recent movers respectively. Data are presented only for those cells in which the correlation was significant at the .10 level for at least one of the three market segments.

### Product Involvement, Product Perceptions and Personal Values

Additional data analyses were conducted to study the pattern of personal value and product evaluation criteria which are associated with specified personal attributes. Table 3 presents the results of these analyses. Since age was noted to be a significant correlate of a variety of personal values related to product perceptions and product use (evidenced in the Table 1 data as well as by Strommen et al., 1972), it was used along with market segment as a classification variable.

As was expected, younger respondents (18 to 34) indicated less dependency and religious conservatism than did older (35 and up) respondents (e.g., see items 5 and 6 in Table 3). The newly moved segment also expressed lesser dependency than the other ("product user") market segments. Interestingly, the results indicate that church leaders, regardless of age, are no more conservative or fundamentalist in orientation than the established nonleader members. If anything, church leaders may be somewhat less conservative than their established followers. This position may enable them to form an important bridge in attempting to attract members from the segment now making product and brand choices (the new movers).

Examinations of product use (based on item 1 as well as actual frequency of reported church attendance) suggests that involvement with the product (in terms of leadership role) is positively related to actual product use and to perceived product need (item 1). In terms of rated product importance, however, involvement in a leadership role does not reflect different perceptions. While church leaders attend more frequently, they report no greater felt importance or meaning for religion (items 2 and 3) than the established members. As expected, newly moved respondents indicated relatively lower usage and importance ratings for the product category. It would, however, seem inappropriate to view this latter segment simply as nonusers and hence of little practical concern. Ratings of the product remain above neutral for this group, no doubt due in part to the fact that these respondents represent newly moved segments who took enough time to fill out a questionnaire on religious values. Thus, there are indications of product interest and at least occasional users in this segment. This may well be a target market for those involved in generating product acceptance.

One set of interesting comparisons is evidenced by the differences in values among younger respondents. In contrast to the earlier conservatism and traditionalism differences, where church leaders were more like the newly moved target segment, younger established members are more like younger movers than are younger leaders in terms of willingness to express negative reactions toward the product (e.g., item 8) and in perceptions of a lack of brand differentiation within the product class (item 10). This may seem to indicate a relatively more "alienated" younger segment exists among church members who are not involved in a leadership role. Yet, concerns for the role of religion in cementing the family (e.g., item 4, plus several others not reported in Table 3) and the general concern with the need for guidance (item 5), suggests that this younger group does have some similar value patterns to the involved leader group(s). In sum, then, the younger, presently uninvolved member may form a kind of psychological bridge to the younger recent mover segment. Use of the former group in attracting the latter segment to product and brand use would seem, however, to be predicated on increasing the nonleaders' actual and felt degree of involvement in institutional decisions.

TABLE 3  
PERSONAL VALUES AND EVALUATION  
CRITERIA ACROSS MARKET SEGMENTS

	Market Segment Means						F Value
	Church Leaders		Church Members		Recent Movers		
	Age		Age		Age		
	<35 (n=27)	>35 (n=78)	<35 (n=41)	>35 (n=130)	<35 (n=115)	>35 (n=80)	
<b>Personal Values:</b>							
1. I feel uncomfortable if I miss going to church each week.	3.30	3.83	2.66	3.50	2.40	2.91	Seg: 16.06** Age: 22.97** Int: 0.69
2. My religion is very important to me.	4.48	4.41	4.08	3.89	3.50	3.77	Seg: 25.83** Age: 2.70 Int: 0.92
3. Without religion, there is little or no meaning in life.	3.96	4.31	3.73	4.35	3.23	3.65	Seg: 13.01** Age: 14.62** Int: 0.47
4. Churches are needed to give children a firm foundation in what's right & wrong.	4.20	4.19	3.83	4.28	3.20	3.70	Seg: 15.92** Age: 7.05** Int: 2.14
5. People on their own are helpless; they need someone (or something) to guide them.	3.50	3.73	3.41	3.91	3.15	3.43	Seg: 4.18* Age: 8.57** Int: 0.52
6. The most important purpose of worship is to find out what God wants us to do.	2.69	3.39	2.78	3.67	2.62	2.87	Seg: 5.37** Age: 25.94** Int: 2.46
7. I feel that my life, at least up till now, has been very worthwhile.	3.85	4.01	4.08	3.89	4.22	3.88	Seg: 0.79 Age: 2.30* Int: 3.59*
8. People in churches are not very friendly to outsiders.	2.04	2.46	2.71	2.58	2.50	2.73	Seg: 6.06** Age: 2.92 Int: 2.38
9. When moving, try to find out about different religions before choosing church.	2.52	2.66	3.10	2.84	2.77	2.55	Seg: 3.63* Age: 0.16 Int: 1.07
10. Most religions are pretty much alike.	2.33	2.21	2.54	2.67	2.79	2.55	Seg: 4.17* Age: 0.43 Int: 0.87
<b>Product Evaluation Criteria:</b>							
1. Social attraction attributes ( $\bar{x}$ of 3 items)	1.81	2.44	2.05	2.23	1.60	2.25	Seg: 0.15** Age: 9.71** Int: 1.01
2. Institutional activity attributes ( $\bar{x}$ of 4 items)	3.16	3.46	2.93	2.56	3.08	3.29	Seg: 1.31** Age: 6.31** Int: 0.09
3. Religious values attributes ( $\bar{x}$ of 2 items)	4.31	4.35	3.88	4.11	3.65	3.90	Seg: 6.83** Age: 2.01 Int: 1.14

\*P<.05; \*\* P<.01

Examinations of product perceptions in terms of the rated importance of evaluation criteria suggest some segment differences. Older respondents, regardless of segment, indicated a greater concern for the church as a gathering place for friends and as a generator of service programs and social activities. Only in the area of religious value criteria did market segment differences appear. Leaders expressed somewhat greater concern than regular members or new movers for the brand as an evaluation criterion and for religious socialization programs for the young. A fair number of areas of agreement appear to exist in terms of criteria deemed important or unimportant in choosing a church. The major difference, consistent with the item 10 results, is a relative lack of felt brand-related differences among those less involved. If brand differentiation is to be a major goal of churches (as it is a concern of church leaders), results suggest that this goal may not be addressing the major choice determinant for nonuser segments. Perhaps a more viable strategy is to address nonspecific product involvement and product use prior to brand loyalty development. The "ecumenical movement" thus appears to relate well to the implications of those market segment differences which seem to exist.

#### Implications

These results are suggestive of differences in activities and information patterns which are associated with what might be termed religious consumer behavior. It appears that product and brand related behaviors do have a real relationship to a variety of consumer attributes and preference patterns. While this hardly seems remarkable to the commercial product marketer, it does indicate the applicability of a basic marketing approach to the social institution marketer involved in encouraging a commitment to "consume" a particular "brand" of services. The decision process of joining a group or of choosing a church can be viewed as a consumer choice, with considerations of product attributes, consumer values and information needs as a crucial component in developing a viable promotional strategy. Despite the relatively intangible quality of social services (such as those provided by a religion), the decision process seems to be structured and determined analogously to more typical consumable product choice phenomena. Examination of this decision process would thus appear to profit from a marketing framework stipulating empirical investigation of decision determinants and behavior correlates. From a theoretic perspective, such extrapolations of approach are important. From a pragmatic standpoint, efficient and effective strategies demand this type of social product application.

#### FOOTNOTES

1. The collection of data was supported by the New England Synod of the Lutheran Church in America.
2. Donald J. Hempel is a Professor of Marketing and Acting Head of the Marketing Department at the University of Connecticut.
3. William J. McEwen is an Assistant Professor in the Communication Division of the University of Connecticut.

4. The home buyer sample was restricted to towns in Massachusetts and Connecticut because population listings were not available for the other four states. Since most of the sampled Lutheran churches were located in these two states (24 of the 30 churches), this restriction had little effect on the comparability of the home buyer sample to the church member sample.
5. Originally, a five page questionnaire was developed. A pretest mailing to 200 Connecticut households was used as the basis for constructing the four-page questionnaire described here.
6. Return rates for the three identified populations were as follows: church members - 29.5%; church leaders - 45.5%; new movers - 21.2%.

#### REFERENCES

- Demerath, N. J. Trends and anti-trends in religious change. In E. B. Shelden and W. E. Moore (Eds), Indicators of social change: Concepts and measurement. New York: Russell Sage Foundation, 1971, 349-445.
- Engel, J. F. Understanding the market for Christianity: A report on Asian Christian strategy. In S. Ward and P. Wright (Eds), Advances in consumer search. Vol. 1. Urbana, Ill.: Association for Consumer Research, 1973. Pp 505-510.
- Kotler, P. A generic concept of marketing. Journal of Marketing, 1972, 36(3), 46-54.
- Kotler, P. & Levy, S. J. Broadening the concept of marketing. Journal of Marketing, 1969, 33(1), 10-15.
- Plummer, J. T. The concept and application of life style segmentation. Journal of Marketing, 1974, 38(1), 33-37.
- Strommen, M. P., Brekke, M. L., Underwager, R. C., & Johnson A. L. A study of generations. Minneapolis: Augsburg Publishing, 1972.
- Tucker, W. T. Future directions in marketing theory. Journal of Marketing, 1974, 38(2), 30-35.

THE COGNITIVE FOUNDATIONS OF  
ATTITUDES: SOME IMPLICATIONS FOR  
MULTI-ATTRIBUTE MODELS

Bobby J. Calder<sup>1</sup>  
University of Pennsylvania

Multi-attribute models of attitude of the type proposed by Fishbein and Rosenberg have received considerable attention in consumer research. Part of the appeal of these models is probably due to their specification as equations. This level of formalization may be more apparent than real, however. As shown in this paper, these models have not been mathematically derived nor have they been induced from empirical relationships. Rather they were originally based on some rather primitive ideas about the cognitive functioning underlying attitudes. They have subsequently become somewhat autonomous from any theoretical justification except the implicit and ill-defined idea that attitudes are based on beliefs. It is argued that in order to improve multi-attribute models it will be necessary to develop a greater understanding of the cognitive foundations underlying them. It is suggested that multi-attribute models are best viewed as "processing rules" and as such may depend on the structural representation of information in the cognitive system. This implies that no single form of a multi-attribute model can be established empirically, and that many forms may be appropriate.

Multi-attribute attitude models have become commonplace in consumer behavior research. They have been applied to attitudes towards products ranging from convenience goods to capital equipment purchases. Although the reason for this popularity is not entirely clear (and might make an interesting study in the sociology of science at that), the usual rationale is that these models provide information about attitudinal structure as well as a basis for behavioral prediction. Knowledge about attitudinal structure, of course, increases the level of our understanding of consumer processes and supplies diagnostic information for decision-making.

Unlike simpler, unidimensional scales, these models are specified in terms of the beliefs about the salient attributes of a product which underlie a person's attitude. Typically the models take a linear compensatory form which may be expressed as  $A_j = \sum_{i=1}^n w_i B_{ij}$ , where  $A_j$  is an overall attitude toward brand  $j$ ,  $B_{ij}$  is a belief about some attribute  $i$  of the brand, and  $w$  is the weight given to belief  $i$ .

Many issues have been actively debated in connection with these multi-attribute models. For one, the variables involved can be conceptualized quite differently (e.g., Bass and Talarzyk, 1972; Cohen, Fishbein, and Ahtola, 1972; Sheth and Talarzyk, 1972). The weighting factor  $w$  may represent the importance of an attribute and the belief factor an evaluation of the brand in terms of that attribute. Or, in a Fishbein-type model, the  $w$  may be taken as the affect

associated with an attribute and the belief factor as the attribute's cognitive expectancy. Other issues center around (1) whether the weights should even appear in the model, (2) whether a disaggregate model in which a regression is performed on all the terms in the model without summation is more appropriate, and (3) whether cross-sectional or individual level of analyses is more suitable (cf. Wilkie and Pessemier, 1973). Finally, even the basic linear compensatory form of the models may be challenged (e.g., Wright, 1972). Instead of a model in which positive and negative attributes cancel each other, conjunctive models in which all attributes must exceed minimum values or disjunctive models in which one or more attributes must exceed some critical value are possible, as well as lexicographic models in which attributes are considered sequentially.

In short, multi-attribute models possess that susceptibility to endless variation and refinement that researchers cannot resist. Plus, they do have predictive ability. But what of their status as compositional models of attitude? In what sense are multi-attribute models related to underlying attitudinal structure? The purpose of this paper is to redirect attention toward such questions and to discuss a few of the issues involved.

#### Background

As correctly noted by Wilkie and Pessemier (1973), "the basic linear compensatory model was developed in social psychology as a static approach to describing an existing attitude structure" (p.437). The most immediate antecedents of the model appear to lie in the early work of Fishbein (1967a) and Rosenberg (1956), as is by now widely recognized. Relatively little attention, however, has been paid to the basis for these original formulations. Fishbein (1967a) initially adapted the perspective of Hullian learning theory. He postulated that an individual first forms a concept, associated with which are a number of responses. These responses constitute a habit-family-hierarchy called a belief system. Beliefs contain mediating evaluative responses which summate through classical conditioning and generalization to become an attitude. In dealing with specific behavioral acts, Fishbein (1967b) adapted the language of Dulany's theory of propositional control as well. Beliefs are thought of as analogous to cognitive expectations or hypotheses about reinforcement contingencies, where the reinforcement possesses some affective value. Rosenberg (1956), in contrast, developed his model in connection with a hypothesized tendency for people to maintain a state of psychological balance between the affective and cognitive components of attitude (cf. Calder, 1973). Multiplying the signs of a cognitive element and an affective element reveals whether the pair is consistent or not. Such thinking evidently led Rosenberg to employ the sum of such products as an index of attitude.

It is apparent that the connection between the original formulations of the Fishbein and Rosenberg models and any underlying attitude theory were at best tenuous, and, in any case, bear no resemblance to current discussions of the linear compensatory model. More recently, it has been argued that the model is basically a subjective expected utility (SEU) approach. The functional similarity of the SEU model and the linear compensatory attitude model does not really provide any theoretical rationale for the attitude model, however. Unlike the attitude model, the maximization of expected utility may be derived from more basic choice axioms. Such thinking does provide a clue as to the implicit thinking behind most investigations of the linear compensatory attitude model.

Much of the appeal of these models is probably due to their specifications as equations. This level of formalization may be more apparent than real. These models have not been mathematically derived nor have they been systematically induced from empirical relationships. Rather they were originally based on some rather primitive ideas about the cognitive functioning underlying attitudes. They have subsequently become somewhat autonomous from any theoretical justification except the implicit and ill-defined idea that attitudes are based on beliefs.

Most investigations of the linear compensatory attitude model are perhaps best characterized as "paramorphic representations." Hoffman (1960) first employed this term in describing the use of linear multiple regression equations as models of human judgment. The equations are intended to predict judgments and to explain some aspects of the judgment process (such as cue weighting), but are not intended to serve as models per se of the underlying psychological process. In the case of the multi-attributes models under consideration, a "paramorphic" research strategy is defensible. But it is a far step from the claims of understanding and diagnosing attitudinal structure which are usually advanced for the models.

#### The Cognitive Foundations of Multi-Attribute Models

It is our contention that an "information-processing" approach will be more fruitful than the typical "paramorphic" approach in further investigating multi-attribute attitude models, whether of the linear compensatory variety or not. The scope of such an information-processing approach was sketched in the early work of Carl Hovland and has been expounded upon by McGuire (1972). Persuasion is considered a multi-stage process involving information (communication) delivery, attention to the information, comprehension, acceptance of the information, persistence of the acceptance, and behavior consistent with the acceptance. Such a framework serves to emphasize the role of cognitive processes as the underlying determinants of attitude.

Of central importance is the role of comprehension. Traditionally comprehension has been identified mostly with rote learning of communication contents or object (product) attributes. It has been argued, however, by Greenwald (1968) and Calder, Insko, and Yandell (1974) that such a view is much too narrow. Comprehension is an active process. Beliefs do not originate only with external sources of information. Individuals also generate new beliefs in processing information and engage existing beliefs stored in memory.

Although cognitive theories of such comprehension are far from complete, any theory must have two crucial aspects: representation and process. Representation refers to the internal information structure; process refers to the rules for manipulating this structure. Cognitive psychologists have recently begun to make a distinction in terms of representation between memory systems and memory codes. Memory systems involve different mechanisms for storage (e.g., long-term versus short-term stores) and their organizational structure. Memory codes refer to the form in which information is stored. A growing number of studies of imagery have investigated visual (iconic) forms of storage. Linguistic codes have continued of course, to receive the most attention. The processing of the information represented has generally been treated in terms of mathematical models, though there has been a trend toward the use of computer simulations (e.g., Anderson and Bower, 1973).



What such considerations have to do with multi-attribute models of attitudes is this. Multi-attribute models are best thought of as putative processing rules. But by themselves such processing rules do not constitute a cognitive theory. It is necessary to specify what it is they operate on, and this specification must be done in both theoretical and operational terms. Let us assume for our purposes that beliefs are represented in the form of linguistic codes. While we have no really adequate theories of the functional units of such codes, it is not unreasonable to assume a propositional format, not unlike the simple declarative sentence (Mehler, 1963; Fillenbaum, 1966; Crothers, 1972). How might these unitary beliefs be organized in memory? The memory system involved may be called active or operational memory. Items in this system are activated either from new input existing in the short-term store or from previously retained material in the long-term store. As is fairly well known, the size of this memory system appears to be constrained to something like six to eight units, though the units themselves may be more or less complex (Miller, 1956). This does not mean that we are limited to considering only seven or so beliefs at a time. It does mean that the organization of beliefs in active memory is usually more complex than a simple list.

According to one hypothesis (Mandler, 1967), a given unit can itself be expanded to seven or so units. The resulting organization is a hierarchical structure of units which are successively redefined at higher levels of organization. Such a pattern is not a list but a list structure (cf. Newell and Simon, 1963; Reitman, 1965). Thus, beliefs relevant to a given topic might be organized so that concepts of increasingly greater specificity are encountered as one moves down through the structure. At the top are very general topics while beliefs at subsequent levels become much less general. We can complete this picture by reconciling these constraints with the seeming ability of people to integrate an endless amount of information in the long run by following Bower (1970) in envisaging metaorganizations of linked hierarchical list structures.

Now consider how the typical multi-attribute attitude model is operationalized. Respondents are given a list of beliefs on a piece of paper which is thought to correspond with those beliefs which are in fact activated by them in considering a given product. There are two things which could happen, depending on whether one takes a pessimistic or an optimistic point of view. Respondents may faithfully search their own belief structure and report on the belief which most closely matches the statement they are presented with. The researcher then applies his own multi-attribute attitude model to this data and derives a prediction of the respondent's attitude. The problem is that this attitude model may or may not make any sense in terms of the belief structure which the person must process in expressing an attitude. It may or may not be a reasonable model of the cognitive processing strategy. Take the case of the typical linear compensatory model. If the person has not had much exposure to a product or time to think about it, so that his belief structure is very simple, perhaps even as simple as the list presented to him, the linear compensatory model may do quite well. This model is an effective way of integrating small, unorganized information structures. If, however, the person has a more complex belief system along the lines of the possibilities we have discussed, only a disjunctive or lexicographic model would be efficient as a cognitive processing strategy. In short, considerable attention must be given to the cognitive representation of beliefs before any processing model can be postulated.

This was the optimistic viewpoint, now for the pessimistic. In giving respondents a list of beliefs to process, we may be presenting them with a very simple problem-solving task. Shortly after working their way through the beliefs,

they are asked to express their attitude. In reflecting on their attitude, it is just possible that instead of processing their own belief structures they simply process the information they have been presented with. The linear compensatory model may be expected to do quite well in this case. Unfortunately, respondents are simply working on a task quite different from the one intended.

### Conclusions

It is our feeling that most applications of multi-attribute attitude models have been arbitrary and indiscriminate. Models have been applied with little consideration for the internal structure being processed or, even worse, the task people are actually performing. The answer to these difficulties lies in two directions. For one, at least as much attention needs to be paid to theoretical accounts of the organization of beliefs as to putative processing models. Further work along the lines initiated by Bettman (1974a,b) in the area of decision net models should prove useful, but there is a great need for innovative work. Second, it is doubtful that simply presenting respondents with lists of beliefs should be further propagated as a research strategy. Considerable effort must be directed toward eliciting beliefs from respondents along with information about how these beliefs are structured in memory.

In closing, it should be noted that we are not arguing that presenting respondents with a list of beliefs is never a defensible research strategy. When the objective, for instance, is to measure the attitude of respondents, endorsement of beliefs as in a Thurstone scaling procedure has considerable merit. The beliefs endorsed are being used as indications of an underlying attitude, just as any other set of behaviors might, with sufficient theoretical justification, be used as an indication of attitude. This is a quite different objective, however, than that of testing a multi-attribute model of attitude.

### FOOTNOTE

1. Bobby J. Calder is Associate Professor of Marketing and Organizational Psychology at The Wharton School, University of Pennsylvania.

### REFERENCES

- Anderson, J. and Bower, G. Human associative memory. Washington, D.C.: Winston, 1973.
- Bass, F. and Talarzyk, W. An attitude model for the study of brand preference. Journal of Marketing Research, 1972, 9, 93-96.
- Bettman, J. Decision net models of buyer information processing and choice: Findings, problems, and prospects. In G. Hughes and M. Ray (Eds.), Buyer/consumer information processing. Chapel Hill: University of North Carolina Press, 1974. (a)
- Bettman, J. Toward a statistics for consumer decision net models. Journal of Consumer Research, 1974, 1, 71-80. (b)
- Bower, G. Organizational factors in memory. Journal of Cognitive Psychology, 1970, 1, 18-46.

- Calder, B. Cognitive consistency and consumer behavior. In H. Kassarian and T. Robertson (Eds.), Perspectives in consumer behavior. Glenview, Ill.: Scott, Foresman, 1973.
- Calder, B., Insko, C. and Yandell, B. The relation of cognitive and memorial processes to persuasion in a simulated jury trial. Journal of Applied Social Psychology, 1974, 4, 62-93.
- Cohen, J., Fishbein, M. and Ahtola, O. The nature and uses of expectancy-value models in consumer attitude research. Journal of Marketing Research, 1972, 9, 456-460.
- Crothers, E. Memory structure and the recall of discourse. In R. Freedle and J. Carroll (Eds.), Language comprehension and the acquisition of knowledge. Washington, D.C.: Winston, 1972.
- Fillenbaum, S. Memory for gist: Some relevant variables. Language and Speech, 1966, 9, 217-227.
- Fishbein, M. A behavior theory approach to the relations between beliefs about an object and the attitude toward the object. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: Wiley, 1967. (a)
- Fishbein, M. Attitudes and the prediction of behavior. In M. Fishbein (Ed.), Readings in Attitude theory and measurement. New York: Wiley, 1967. (b)
- Greenwald, A. Cognitive learning, cognitive response to persuasion, and attitude change. In A. Greenwald, T. Ostrom and T. Brock (Eds.), Psychological foundations of attitudes. New York: Academic Press, 1968.
- Hoffman, P. The paramorphic representation of clinical judgment. Psychological Bulletin, 1960, 57, 116-131.
- Mandler, G. Organization and memory. In K. Spence and J. Spence (Eds.), The psychology of learning and motivation, I. New York: Academic Press, 1967.
- McGuire, W. Attitude change: The information-processing paradigm. In C. McClintock (Ed.), Experimental social psychology. New York: Holt, 1972.
- Mehler, J. Some effects of grammatical transformations on the recall of English sentences. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 346-351.
- Miller G. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.
- Newell, A. and Simon, H. Computers in psychology. In R. Luce, R. Bush, and E. Galanter (Eds.), Handbook of mathematical psychology, I. New York: Wiley, 1963.
- Reitman, W. Cognition and thought: An information-processing analysis. New York: Wiley, 1965.
- Rosenberg, M. Cognitive structure and attitudinal affect. Journal of abnormal and social psychology, 1956, 53, 368-372.
- Sheth, J. and Talarzyk, W. Perceived instrumentality and value importance as determinants of attitudes. Journal of Marketing Research, 1972, 9, 465-467.

Wilkie, W. and Pessemier, E. Issues in marketing's use of multi-attribute attitude models. Journal of Marketing Research, 1973, 10, 428-441.

Wright, P. Consumer judgment strategies: Beyond the compensatory assumption. In M. Venkatesan (Ed.), Proceedings of the third annual conference of the Association for Consumer Research, 1972.



THE PROCESS OF ATTITUDE ACQUISITION: THE VALUE OF A DEVELOPMENTAL  
 APPROACH TO CONSUMER ATTITUDE RESEARCH<sup>1</sup>

Jerry C. Olson and Andrew A. Mitchell<sup>2</sup>

Pennsylvania State University

Given the current controversies in attitude and communications research, it is suggested that substantial effort be focused on the process of attitude acquisition which is the basis for the phenomena of attitude structure and change. A model of the attitude acquisition process based on behavioral learning theory mechanisms is presented. The authors suggest that any attitude model (e.g., expectancy-value models) should be described in terms of the developmental processes involved in acquiring the model components. An example is presented of the developmental aspects underlying the Fishbein attitude model. The implications of this developmental perspective for attitude and communication research are discussed.

Over the last 40 or 50 years in psychology and at least over the last decade in marketing, immense efforts have been devoted to the study of attitudes and attitude-related phenomena such as attitude change. In particular, consumer researchers, beginning perhaps with Hansen (1969), have become enamoured with attitudes, especially with attitude models of the expectancy-value variety.

Consumer behaviorists have used the concept of attitude for a variety of applied purposes, ranging from predicting consumer choice behavior (e.g., Bass & Talarzyk, 1972; Ginter, 1974), to serving as a basis for market segmentation (e.g., Mitchell, 1974; Myers & Alpert, 1968), to use as a criterion for advertising effectiveness (e.g., Lavidge & Steiner, 1961; Greene & Stock, 1966; Winter, 1973), in addition to its ubiquitous inclusion as an explanatory construct in theoretical models of consumer behavior (cf. Engel, Kollat, and Blackwell, 1973; Hansen, 1972; Howard & Sheth, 1969; Nicosia, 1966; Sheth & Park, 1973).

Given the widespread use of the attitude construct in consumer behavior research, it is surprising to find few (if any) explicit discussions of the antecedent conditions and dynamic processes underlying attitude formation. On the contrary, much attitude research is cross-sectional and treats attitude as given, a variable to be measured and related to other concepts such as behavioral intention, purchase, or information search. Those studies which have examined the attitude construct in an experimental and/or longitudinal fashion generally have not explicitly discussed the factors and processes which cause attitudes, but rather seem to suggest that attitudes rather "magically" occur upon, for example, exposure to an advertising communication or after some purchase/use experience.

The above-stated comments are not intended to suggest that previous research is without value. Clearly, some operational and conceptual advances have been made in attitude research (e.g., see Day, 1973; Wilkie & Pessemier, 1973). However, we do intend to suggest that ignoring or side-stepping the basic conceptual issues underlying attitude acquisition and the related topic of atti-

tude structure generally has been dysfunctional and has led to several problems which are clearly evident in the consumer research literature, particularly in the areas of communications/advertising research and attitude models.

For example, research concerning the effect of communication on consumer attitudes and behavior seems to have been strongly influenced by the methodology and descriptive problem orientation of Hovland, Janis, and Kelley (1953). The typical study in this "Yale approach" can be described by the theme sentence, "Who says what to whom with what effect?". In most studies the dependent variable, "with what effect", was a self-reported measure of behavior or attitude change. The independent variable usually involved either the communication source ("who"), the structure or content of the communication ("what"), or certain characteristics of the communication recipient ("to whom"). Partially because of this atheoretical influence, few studies of communication/advertising effectiveness or attitude change have attempted to examine the process of creating or changing attitudes and the factors involved therein. (See Wright [1973] for a notable exception, although he does not specifically discuss attitude formation.)

In our opinion, this research approach has hindered the development of a real understanding of communication and advertising effects. For instance, much advertising today seems to communicate in a nonverbal manner. For example, many so-called image advertisements merely present a brand against a pictorial background--e.g., Marlboro or Salem cigarette ads--without any direct persuasive copy. Yet, communication and advertising researchers, to the best of our knowledge, have not evolved a theory that will explain such phenomena, and in fact have studiously ignored this topic.

Since attitude has long been a primary criterion measure of advertising effectiveness, we believe that a thorough conceptual examination of the processes and factors involved in attitude acquisition and change may provide insights regarding the effects of "image advertisements" as well as the typically studied, more directly persuasive communications. Certainly, an understanding of the processes involved in communication effectiveness would be useful to marketing management and certain regulatory agencies.

The dysfunctional effects of ignoring such attitude acquisition phenomena are perhaps even more obvious in the burgeoning research literature on expectancy-value attitude models. This research contains numerous examples of non-consistent use of terminology, imprecise conceptualizations, and miscellaneous controversies and confusions (e.g., witness the exchanges between Cohen, Fishbein, and Ahtola, 1972; Bass, 1972; Sheth, 1972, and Talarzyk, 1972). Often consumer researchers seem to treat the components of expectancy-value attitude models as given, frequently modifying them for operational ease, without considering the conceptual and theoretical logic upon which the models are based. A recent review (Wilkie & Pessimier, 1973) indicates the extent to which such a non-theoretical orientation results in a field becoming "bogged down" in predictive operational issues (e.g., should one include an "importance" score?), while ignoring the critically important theoretical basis for the model. Consequently, seldom if ever in the marketing literature have the basic theoretical concepts underlying attitude acquisition and change (and therefore the models themselves) been put to a rigorous test. Thus, the present massive body of attitude model data is built upon a weak, unvalidated conceptual base.

To summarize thus far, we contend that consumer researchers have ignored a critical phase in their research on attitudes and now need to "back-track" to that point and conduct several important studies. Specifically, we believe that consumer behaviorists should devote substantial research attention to the

process by which attitudes are formed or acquired, a process which underlies the important phenomena of attitude structure and attitude change. Thus, when attitude models are used in our research, we should identify and validate the processes by which the components of that model are acquired by the consumer. Even if empirical validations are not conducted, we should at least force ourselves to logically describe in conceptual terms the acquisition processes involved in our attitude models.

Such theoretical and empirical attention to the attitude acquisition process could have a variety of benefits. For example, by forcing researchers to be more precise in their conceptual thinking, confusing and conflicting terminology would be clarified. Thus, our hypothetical concepts would be developed to fit the underlying theory rather than the researcher's problem area or purpose.

Another benefit derived from concentrating upon the attitude acquisition process would be to aid and encourage the formulation of an explicit, clear attitude construct and of a formal conceptualization of attitude structure. Such theoretical developments would in turn encourage researchers to validate these notions, the success of which would provide investigators with a firmer conceptual base.

Moreover, an explanation of attitude formation would be useful in developing an explanation of the attitude change process. With an explicit conceptualization of how attitudes change, communications research would be on firmer theoretical ground in the past, and perhaps findings of a generalizable nature could at last be developed. Also, with a theoretical base from which to work, advertising researchers may be more willing to study difficult phenomena such as image advertising or nonverbal communication.

Finally, an explication of the process of attitude acquisition would provide a firmer and less ambiguous relationship between information processing phenomena and attitude structure. Indeed, it seems clear that attitude development occurs through information processing. What researchers need now is a description, in theoretical terms, of the process by which one incorporates information and acquires an attitude.

In order to stimulate research and hopefully begin to achieve some of the above-stated benefits, the following section presents a conceptual model of the attitude acquisition process based upon behavioristic learning theory. This model was chosen for its parsimony, the precision of its concepts, the amount of past research it has generated (especially in the psychology literature), and simply because it is the only model of which we are aware. Although we have made several minor conceptual modifications to this model of attitude acquisition, essentially the following discussion is a summary of the work of others (as referenced).

#### The Process of Attitude Development, Formation, or Acquisition

Although seemingly ignored by most marketing researchers and social psychologists, a substantial body of conceptual and empirical work in psychology has addressed the issue of attitude<sup>3</sup> acquisition in terms more specific than the typical lip-service that "attitudes are learned through experience." The usual theoretical perspective of this research is drawn from behavioristic learning theory<sup>4</sup> and, consequently, relies heavily upon the basic learning mechanisms of classical and instrumental or operant conditioning (as well as the more complex mechanisms of generalization) for explanations of the attitudinal acquisition process.



Doob (1947), who was perhaps the first to apply learning theory principles to the attitude formation process, considered attitude to be ". . . (1) an implicit response, (2) which is both (a) anticipatory and (b) mediating in reference to patterns of overt responses, (3) which is evoked (a) by a variety of stimulus patterns (b) as a result of previous learning or of gradients of generalization and discrimination, (4) which is itself cue- and drive-producing, (5) and which is considered socially significant in the individual's society." Generally, later theorists maintained the essence of Doob's conceptualization while frequently emphasizing, after Allport (1935), that the implicit (i.e., internal) attitudinal response is of an evaluative nature (cf. Rhine, 1958; Staats & Staats, 1958). Our interest, of course, is with the process by which a subject acquires that implicit evaluative response called attitude.

Before examining the specific learning processes by which one acquires a particular attitude toward a stimulus object, it should be noted that the conceptual perspective of the present paper is grounded in the theoretical positions of Osgood (e.g., Osgood, Suci, & Tannenbaum, 1957) and Fishbein (cf. 1967), among others, in that all stimulus objects are considered to have associated with them two implicit, mediating responses (see Figure 1). That is, a stimulus object elicits two internal responses from a subject. One of these mediating responses represents the identification or categorization of the stimulus object and is often termed a concept-labeling response, while the other internal response is the attitudinal, affective or evaluative response. This attitudinal response may be positive or negative in varying degrees of intensity or, in the case of

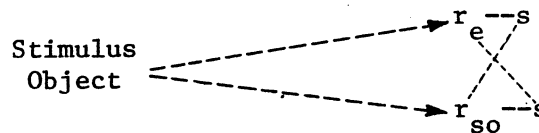


Figure 1. A stimulus object elicits two implicit responses, an evaluative response with associated stimulus properties ( $r--s$ ) and a labeling response with cue properties ( $r^e--s$ ), both of which become conditioned to, and tend to elicit, one another.

novel, unfamiliar stimuli, the evaluative response may be neutral. The critical aspect of this theoretical notion is that the elicitation of the evaluative attitude response (either positive, negative, or neutral) is "automatic" upon exposure to and identification/categorization of the stimulus object (see Fishbein, 1967, for a more detailed presentation and justification of this position).

To summarize, note in Figure 1 that a stimulus object elicits two internal, mediating responses, one a concept-labeling response and the other an evaluative, attitude response. Furthermore, over time, through classical conditioning, the labeling and evaluative responses become conditioned to, and are elicited by, one another. Note, however, that this theoretical position merely states that persons have attitudes (in the form of evaluative responses) to all discriminable stimuli, but does not account for the acquisition of a specific attitude -- i.e., one with a particular degree of positive or negative direction. Before discussing that acquisition process, let us first examine the processes by which one may acquire the concept-labeling response to a stimulus object.

#### Concept Formation Through Direct Conditioning

Before an attitudinal evaluative response to a stimulus object can be acquired, one must first be able to discriminate, identify, and categorize that stimulus. These latter abilities are developed as one comes to recognize that a set of specific stimuli "belong together" or form a discriminable stimulus

pattern, and thus are given a specific meaning<sup>5</sup>, typically in the form of an internal labeling response. Such learning has been termed concept formation (cf. Rhine, 1958). For instance, in our culture a stimulus object with four legs, a back, and a seat is usually recognized as a chair while a four-legged, backless object with a seat is commonly labeled a stool. The process of learning such concepts (i.e., attributing meaning to stimuli) takes place over a period of time and is a function of both classical conditioning and reinforcement contingencies. For instance, parents may continually point out and verbally identify objects for their children and, moreover, may provide positive reinforcement when the child demonstrates concept learning by making the correct overt, verbal labeling response. Thus, as depicted in Figure 2, first-order concept formation may be conceptualized as the acquisition of an implicit response to a stimulus or set of stimuli (cf. Osgood, 1952). The response is typically considered to be an internal, labeling response with associated cue (stimulus) properties which may in turn mediate (or guide) other internal or overt responses.

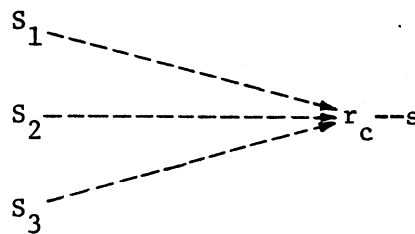


Figure 2. First-order concept formation occurs when a specific pattern of stimuli ( $S_1$ ,  $S_2$ , and  $S_3$ ) come to elicit, through classical and/or instrumental conditioning, an implicit labeling response ( $r_c$ ) which possesses certain cue properties ( $s$ ), and which may mediate other responses.

#### Concept Formation Through Mediated Generalization

It is perhaps more realistic conceptually, especially for multi-attribute brands and other complex marketing stimuli, to discuss a somewhat more complicated concept formation process based upon the principle of mediated or secondary generalization (e.g., Cofer & Foley, 1942; Eisman, 1955). This more involved type of concept acquisition is essentially analogous to direct conditioning. However, instead of external, physical stimuli coming to elicit an internal, labeling response, in this case the concept labeling response is a generalization of previously learned, internal, concept responses. Since the generalized, higher-order labeling response is a function of other internal, mediating responses, the process is termed secondary or mediated generalization.

To illustrate secondary generalization of a concept labeling response, consider in Figure 3 a relatively complex stimulus object such as a product brand "X" which possesses three discriminable attributes A, B, and C (e.g., price of \$4.98, red color, large package size). Each of these three attributes is, in fact, an array of stimuli which, through previous conditioning, has come to elicit its own concept labeling response and, of course, its own implicit evaluative response (see Figure 2). If the three attributes A, B, and C are frequently experienced together, i.e., contiguously as they would be if they were stable characteristics of a brand, the three labeling responses and their respective stimuli (i.e.,  $r_a--s$ ,  $r_b--s$ , and  $r_c--s$ ) may come to be associated with and thus elicit a<sup>a</sup>secondary, general<sup>b</sup>ized, internal labeling response ( $r_x--s$ ) which represents the meaning of the overall stimulus object,

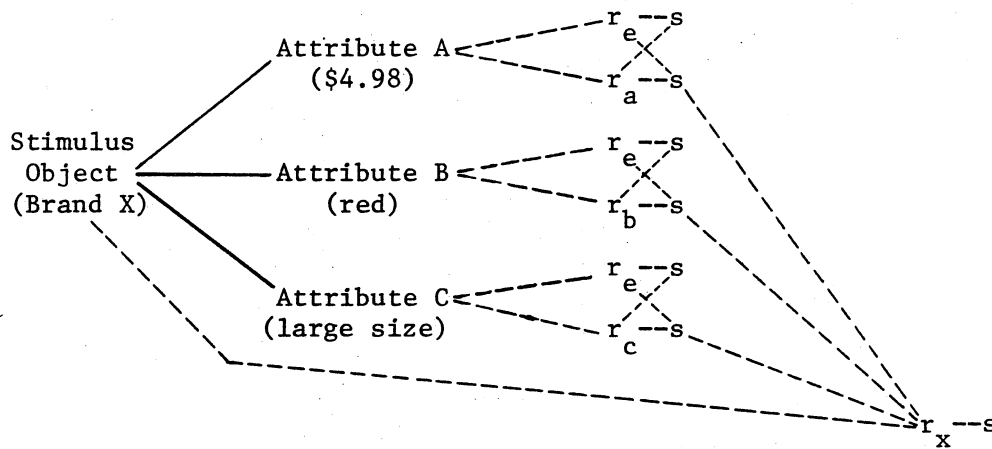


Figure 3. Formation of a second-order concept response ( $r_x--s$ ) to a stimulus object by secondary, mediated generalization of the set of first-order labeling responses ( $r_a--s$ , etc.) associated with the attributes of the stimulus object (X).

brand X. Usually this internal labeling response will be the brand name. Of course, concept formation through mediated generalization is facilitated by both classical conditioning (frequent exposure to brand X) and reinforcement contingencies (rewards or punishments for overt, verbal labeling responses or other overt behaviors).

#### Acquisition of the Evaluative, Attitudinal Response

Now let us turn our attention to the process by which one acquires a specific attitude response to a stimulus object (i.e., an evaluative response with a particular degree of positive or negative intensity). The attitude acquisition process may occur in essentially two ways: (a) through direct conditioning of the evaluative response by classical or reinforcement conditioning mechanisms, or (b) through more indirect conditioning of the evaluation response by mediated generalization mechanisms, in addition to classical and instrumental conditioning influences.

Attitude Acquisition Through Direct Conditioning. A particular evaluative response to a stimulus object may be acquired directly through classical conditioning mechanisms (cf. Staats & Staats, 1958). For example, a positive attitude toward brand "X" may be established by repeatedly and contiguously presenting a subject with brand "X" and another stimulus which elicits the desired evaluative response either innately (e.g., food) or through previous conditioning (e.g., colorful sunsets). Conversely, a negative evaluative response could be conditioned to brand "X" by repeatedly pairing it with a stimulus which elicits the desired negative response (e.g., electric shock [Bandura & Rosenthal, 1966]).

Alternatively, attitude responses may also be acquired through reinforcement mechanisms. For example, by rewarding overt, verbal evaluations of a positive nature (e.g., Eisman, 1955), a positive, internal, evaluative response may be conditioned.

Attitude Acquisition Through Mediated Generalization. For multi-attribute stimuli of major interest to consumer researchers, the usual process by which a particular evaluative response is acquired may be considered to be "higher-order" conditioning based upon the principle of mediated generalization, in addition to classical and operant conditioning mechanisms. To illustrate that

process, consider the acquisition of a specific attitude (of a particular direction and intensity) towards a product brand, X (see Figure 4).

Suppose that brand X has three attributes or characteristics (A, B, and C) associated with it, each of which elicits its own unique, previously learned labeling and evaluative responses (e.g.,  $r_{e--s}^a$  and  $r_{e--s}^c$ ). The reader may recall from the discussion above, that a second-order, concept labeling response

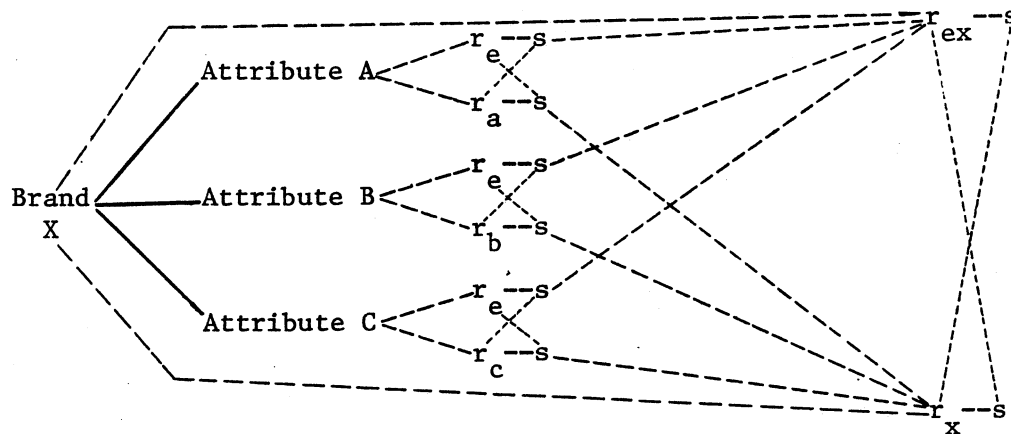


Figure 4. Brand X has three characteristics or attributes (A, B, and C), each with previously-acquired labeling and evaluative responses. The overall evaluative response to brand X ( $r_{e--s}^{ex}$ ) is a function, through generalization, of the combination of the separate evaluative responses to the attributes associated with brand X. Moreover, through exposure to and experiences with brand X, (i.e., classical and instrumental conditioning, respectively), brand X may come to elicit the generalized overall evaluative and labeling responses directly.

( $r_{e--s}^x$ ) to brand X is acquired through the process of secondary, mediated generalization (and perhaps also classical and reinforcement conditioning). By a similar process of mediated generalization, an overall evaluative response may be acquired towards brand X. That is, the evaluative attitudinal response to brand X ( $r_{e--s}^{ex}$ ) is generalized from, or derived from, the various mediating evaluative responses associated with the attributes comprising brand X (see Figure 4). Thus, through the mechanisms of mediated generalization, and perhaps classical and reinforcement conditioning, as well, a consumer acquires a particular evaluative response, an attitude, towards brand X.

The notion that the overall, evaluative, attitudinal response to a stimulus object is a generalization from specific evaluative responses to the attributes (or concepts) which are associated with that stimulus object is a basic theoretical position of many attitude researchers (e.g., Osgood & Tannenbaum, 1955; Peak, 1955; Rosenberg, 1956; Fishbein, 1963), although some do not explicitly recognize this basic assumption (e.g., Thurstone, 1928). Not only does this notion make good logical sense, but it has also received indirect empirical validation (e.g., Fishbein and Hunter, 1964), although the specific form of the mediated generalization (i.e., the combination of evaluative responses) is not agreed upon (see Anderson & Fishbein, 1965; for a "test" of two alternative combination forms).

However, irrespective of the generalization/combination procedure favored (additive, averaging, differentially weighted, etc.), we believe that it is

extremely important, conceptually and operationally, to remind ourselves of the basic theoretical premise that the formation of a specific attitude toward a complex stimulus object is a function of the combined evaluative responses to the attributes of that stimulus object.

#### Evidence for Validity of the Attitude Acquisition Process

There is considerable evidence in the psychological literature which supports the attitude acquisition process described above. Staats and Staats demonstrated in several studies that attitudes may be acquired through higher-ordered classical conditioning, that is, by pairing the attitude object with a stimulus previously conditioned to elicit a particular evaluative response (Staats & Staats, 1957, 1958, 1959; Staats, Staats, & Biggs, 1958; see also Staats, 1967). Furthermore, Staats, Staats, & Crawford (1962) successfully demonstrated first-order classical conditioning of a negative attitude by pairing words (the CS) with shocks or loud noises (the UCS), as has Maltzman (1968). Evidence also exists to support an instrumental conditioning explanation of attitude acquisition (e.g., Eisman, 1955; Scott, 1957, 1958), although Scott's results have been questioned (Dahlke, 1963; Kiesler, 1965).

It must be noted that the experimental designs used in the typical attitude conditioning study have been criticized as containing inherent demand characteristics (Orne, 1962) which can provide an alternative explanation for the results (cf. Kiesler, Collins, & Miller, 1969). However, recent work has minimized the demand characteristics of the experimental procedures and has still obtained a conditioned evaluative response (e.g., Zanna, Kiesler, & Pilkonis, 1970).

Finally, the impression formation literature (see review by Tagiuri, 1968) provides additional evidence to support the attitude acquisition process presented above. In the typical experiment, subjects receive information--often in the form of adjective traits (which elicit certain positive or negative evaluative responses)--about a hypothetical person (initially neutrally evaluated) and then record their impressions (evaluation) of that person. As Chambers (1969) has noted, impression formation is very similar to higher-ordered classical conditioning and has yielded similarly successful results (cf. Anderson & Hubert, 1963). Moreover, research in impression formation has also examined repetition and order effects, variables generally ignored by the neobehaviorist attitude researchers but factors highly relevant to advertising (e.g., Chambers, 1971; Shapiro and Tagiuri, 1958).

In summary, the above-cited studies indicate that specific attitudes can be acquired and changed through the mechanisms of classical and operant conditioning and mediated generalization. Thus, the basic principles of the conceptual attitude acquisition model presented above have been shown to have some validity.

#### Applications and Implications of a Process Approach to Attitude Development

In the absence of an alternative theoretical explanation of the process of attitude acquisition which possesses the precision and comprehensiveness of the learning approach presented above, we suggest that this behavioristic theory be used--very explicitly--in future attitude and communication research by consumer behaviorists. The remainder of this paper briefly describes some of the ways in which an explicit consideration of the attitude acquisition process can improve consumer attitude research.

## Development and Validation of Attitude Models

It is, of course, desirable that consumer researchers continually evolve new models and develop modifications of existing models of attitude structure. However, we contend that scientific progress in our discipline is retarded if such "newcomers" are not subjected to certain standards. As we suggested at the beginning of this paper, a minimal "standard" should be a conceptual (and logical) description of how the components of a proposed model could have developed and been acquired by a consumer. Moreover, given the theoretical perspective on attitude acquisition presented here and elsewhere (cf. Fishbein, 1967), one might expect developers of new models to compare and contrast their attitude models with earlier work. Ideally, such inter-model comparisons would involve not only empirical (predictive) differences, but also would identify conceptual distinctions.

A comparative orientation to attitude model research will, in our opinion, limit the current proliferation of attitude models to those whose components and developmental processes are clearly different--conceptually--from those presented here. Rigorous comparison may reveal several currently-used models to be merely the "old" model with new or modified terminology, but essentially the same concepts. Moreover, a rigorous comparison tradition may find that certain "attitude models" are not actually models of attitude, but rather are descriptions of some other construct. The whole process of positioning one's hypothetical constructs vis-a-vis previously postulated constructs is a scientifically healthy practice and could be applied profitably to areas of consumer research other than attitude formation.

An example. One currently popular and occasionally controversial attitude model which can stand up well under the scrutiny advocated above is that of Fishbein (e.g., 1961, 1963), a basic expectancy-value attitude model similar to those of other researchers (Peak, 1955; Rosenberg, 1956; Vroom, 1964). Fishbein's attitude model is, in fact, completely consistent with the acquisition process presented above, since its development seems to have been influenced by the neo-behavioristic, learning theory approach to attitude formation (e.g., Staats & Staats, 1957, 1958; Rhine, 1958). Since Fishbein has presented the developmental processes underlying his model quite clearly (especially see Fishbein, 1967) and because most consumer researchers are familiar with that model, the following only summarizes his rationale.

Fishbein (cf. 1967) considered attitude to be a unidimensional evaluation, thereby removing the cognitive dimension from the attitude construct. Instead, he conceptualized a belief construct which is separate from, but related to, the evaluative concept and which incorporates cognitive aspects of the stimulus object. To Fishbein, a belief is the subjective perception of association between the belief object and some other object, attribute, or concept, and thus beliefs are often operationalized as the rated "probability" or "improbability" that a relationship exists between the object of belief and the other concept (Fishbein & Raven, 1962).

The notion of belief fits into the attitude acquisition process in the following way. Given that the belief object is a stimulus and related objects, attributes, or concepts are represented psychologically as internal, mediating, labeling responses (see Figure 3), a belief may be considered as an S-r association. From this perspective a belief will be stronger (i.e., the subjective probability of relationship will increase) as the S-r association becomes stronger between a belief object and its related attributes; for example, as a function of the number of contiguous S-r experiences or with an increase in the magnitude or number of rewarding stimuli following a S-r association.

Clearly, the beliefs one has regarding a stimulus object may be hierarchically arranged according to strength to form a belief-family-hierarchy (see Fishbein, 1963). Thus, there are at least two ways of operationalizing strength of beliefs: (a) self-reported probability of association between the belief object and another concept, and (b) order in the belief-family-hierarchy (perhaps measured by order of free association responses). It should be emphasized that Fishbein includes only "salient" beliefs in his attitude model--i.e., those beliefs high in the hierarchy, the strongest beliefs.

Consistent with the conceptualization of the attitude formation process discussed in the present paper, Fishbein's model specifies that the overall evaluative, attitudinal response to a stimulus object is a function of the various evaluative responses to the concepts or attributes believed associated with the object. Moreover, the Fishbein model explicitly states that the extent to which single evaluative responses to attributes of a stimulus object affect the overall attitude toward the stimulus object is weighted by the strength of the belief regarding the relationship between the stimulus object and that attribute. Fishbein precisely specifies the form of this mediated generalization process by proposing that the overall evaluative response to a stimulus object is determined by summing the product of the strength of belief and the intensity of the evaluative response to each attribute associated with the stimulus object. Mathematically, his model is:

$$A_o = \sum_{i=1}^n B_i a_i$$

where:  $A_o$  = the attitude toward object  $o$

$B_i$  = the strength of belief  $i$  about  $o$ , that is, the "probability" or "improbability" that  $o$  is associated with some concept  $x_i$

$a_i$  = the evaluative aspect of  $B_i$ , that is, the evaluation of  $x_i$

$n$  = the number of beliefs about  $o$ , that is, the number of salient beliefs in the individual's belief-family-hierarchy

To our knowledge, no other attitude model has been so clearly described in terms of the processes involved in acquiring the components of the model. We reiterate our earlier suggestion: An explicit conceptualization of the attitude acquisition process would be useful, especially heuristically, for all attitude researchers.

### Implications for Attitude Modeling

As one illustration of the heuristic implications for attitude modeling of a developmental perspective, note that an explicit theoretical explanation for the acquisition of attitude model components clearly reveals the origins and causal factors underlying the model. From such a perspective it is relatively easy to develop experimental manipulations of model components and thus create experimental tests of the model. For example, it is now obvious how both the belief and evaluative-aspect-of-belief components in Fishbein's model might be manipulated in an experimental test of the model's predictive power. Such a perspective is obvious only because we now have a conceptual paradigm which explains the developmental processes underlying the model.

### Implications for Communications Research

When one considers an expectancy-value model such as Fishbein's from a developmental process perspective, certain interesting and theoretically important communication phenomena become more obvious. In fact, combining the developmental processes and the basic Fishbein model, for example, may form the basis for a theoretical explanation of how communications affect attitudes. For instance, exposure(s) to a communication which simply presents a stimulus object-attribute pair could create (through classical conditioning) a belief that the particular attribute is associated with the stimulus object. As this belief becomes stronger the evaluative aspect of the belief would have an increasing influence on the overall attitude toward the stimulus object. Thus, conditioning this belief to the stimulus object, through communication messages in this case, would cause an increase or decrease in the favorability of the overall attitude toward the object (depending upon the evaluative aspect of the belief). Note that a similar change in overall attitudinal favorability could be affected by increasing or decreasing the evaluative aspect of an already-conditioned belief. Research on such phenomena would not only benefit consumer behaviorists interested in attitude change phenomena, but should also stimulate communications researchers who seem to lack specific criterion measures for the effects of their independent variables.

Moreover, the attitude acquisition process presented above encourages the examination of communications such as image advertising. For example, we know that if two stimuli are presented contiguously over a period of time one will acquire an S-R association between the two, and we know that a belief is a type of verbalized S-R association. Therefore, even non-verbal, image advertising may create beliefs. Such beliefs can be measured and the effect of their acquisition on overall attitude can be empirically validated, given the concepts and operational procedures specified by Fishbein, for example (cf. Fishbein & Raven, 1964).

### Need for Longitudinal Research Designs

Another consideration prompted by the developmental perspective to attitude acquisition, is the clear need for longitudinal research designs. In order to test many of the notions discussed above, most of which are entirely within-subject phenomena--that is, the processes occur within an individual subject--we need multiple measures of an individual's beliefs and the evaluative aspects of those beliefs at several points in time, as a function, for example, of exposure to a series of communications. Only with a longitudinal design can changes in the determinants of attitude and subsequent attitude change be adequately inferred.

Finally, it is of distinct interest to note that the data from a set of experimental manipulations of model components, in a longitudinal within-Ss design, would provide a more direct test of attitude models than the cross-sectional studies typically conducted. Moreover, if manipulations of the components of competing models (e.g., ideal-point model vs. the Fishbein model) be included in the same study, one would have the makings of a critical test (cf. Platt, 1964) which could provide relatively unambiguous information regarding the validity and predictive value of each model. Hopefully, a program of such comparative research would provide a better understanding of attitude change processes and influences as well as a set of "better-validated" models of attitude structure. The present authors are currently conducting such a programmatic series of studies.



## Summary

In this paper, we have argued for an examination of the processes underlying acquisition of a specific attitude. The model presented and advocated here is derived from the behavioristic learning theory approach to attitude formation (see Doob, 1947; Fishbein, 1967; Staats, 1967). This model makes the following basic points. (a) The hypothetical construct of attitude is most usefully considered as unidimensional evaluation or affect. (b) Any stimulus object automatically elicits two implicit responses, (1) a concept labeling response and (2) an evaluative response, both of which possess stimulus properties capable of eliciting other implicit or overt responses. (c) These implicit responses may be acquired either through the mechanisms of classical conditioning, instrumental conditioning, or mediated generalization, or some combination thereof.

We believe that a close examination of the attitude construct from a development process perspective, such as that provided by the present model, will enrich consumer attitude research, in the following ways. (a) Researchers will be forced to be more precise, conceptually and operationally, in their use of attitude and attitude-related constructs. (b) The process approach will encourage the development of models of attitude structure, such as expectancy-value models, which are theoretically consistent with these acquisition concepts (or with some other conceptualization of attitude formation). (c) Theoretically-based attitude models encourage researchers to more directly test the basic notions (now explicitly stated) which underlie the model, thus yielding validated frameworks. (d) Attitude models which are based upon a validated developmental process may be more easily used to structure and predict a variety of attitude-related phenomena (e.g., attitude change, non-verbal communication effects). Interesting, heuristic perspectives seem to open up when one bases his research on a theoretically sound conceptual framework.

## FOOTNOTES

1. Preparation of this paper was partially supported by a grant to the first author from the Center for Research, College of Business Administration, Pennsylvania State University.
2. Both authors are Assistant Professors of Marketing, The Pennsylvania State University.
3. In this paper, attitude is considered to be a unidimensional evaluative construct, equivalent to affect (liking) towards, or an evaluative judgment (good-bad) of, a stimulus object. This view is consistent with most of the behavioristically oriented researchers who are concerned with attitude formation, for example, Doob (1947), Eisman (1956), Staats & Staats (1958), and Fishbein (1967). However, it should be noted that the tripartate conceptualization of attitude, with affective, cognitive, and conative components, has numerous apostles and certain unique advantages (cf. Triandis, 1971) as well as disadvantages, particularly operational problems (Fishbein, 1967).
4. Van De Geer and Jaspars (1966) have termed this theoretical perspective the neobehavioristic approach to cognition.
5. Note that this process of assigning meaning to stimulus patterns is essentially equivalent to the formation of chunks in information processing (cf. Newell & Simon, 1972; Olson, 1974, p. 34).

6. The processes by which attributes become associated with or are linked to a stimulus object are of enormous interest to consumer researchers and are discussed later in this paper. To give only two brief examples, however, the associations between the attributes and brand X (N.B., conceptualized as beliefs by Fishbein, 1963) may be formed by either presenting contiguous brand X-attribute combinations through advertising communications, or through the "natural" reinforcement and contiguity mechanisms inherent in product usage experiences.

## REFERENCES

- Allport, G. W. Attitudes. In C. Murchison (Ed.), Handbook of social psychology. Worcester, Mass.: Clark University Press, 1935, 798-884.
- Anderson, L. R., & Fishbein, M. Prediction of attitude from the number, strength, and evaluative aspect of beliefs about the attitude object: A comparison of summation and congruity theories. Journal of Personality and Social Psychology, 1965, 3, 437-443.
- Anderson, N. H., & Hulbert, S. Effects of concomitant verbal recall on order effects in personality impression formation. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 379-391.
- Bandura, A., & Rosenthal, T. L. Vicarious classical conditioning as a function of arousal level. Journal of Personality and Social Psychology, 1966, 3, 54-62.
- Bass, F. M. Fishbein and brand preference: A reply. Journal of Marketing Research, 1972, 9, 461.
- Bass, F. M., & Talarzyk, W. W. An attitude model for the study of brand preference. Journal of Marketing Research, 1972, 9, 63-72.
- Chambers, D. K. Repetition and order effects in attitude formation. Journal of Personality and Social Psychology, 1971, 17, 219-228.
- Cofer, C. N., & Foley, J. P. Mediated generalization and the interpretation of verbal behavior: I. Prolegomena. Psychological Review, 1942, 49, 513-540.
- Cohen, J. B., Fishbein, M., & Ahtola, O. T. The nature and uses of expectancy-value models in consumer attitude research. Journal of Marketing Research, 1972, 9, 456-460.
- Dahlke, A. E. The effects of reinforcement and punishment on attitude change. Unpublished doctoral dissertation, University of Minnesota, 1963.
- Day, G. S. Theories of attitude change and structure. In S. Ward & T. Robertson (Eds.), Consumer behavior: Theoretical sources. Englewood Cliffs, N. J.: Prentice-Hall, 1973.
- Doob, L. W. The behavior of attitudes. Psychological Review, 1947, 54, 135-156.
- Eisman, B. S. Attitude formation: The development of a color preference response through mediated generalization. Journal of Abnormal and Social Psychology, 1955, 50, 321-326.

- Engel, J. F., Kollat, D. T., & Blackwell, R. D. Consumer behavior. (2nd edition) New York: Holt, Rinehart, & Winston, 1973.
- Fishbein, M. A theoretical and empirical investigation of the interrelationship between beliefs about an object and the attitude toward that object. Unpublished doctoral dissertation, University of California, Los Angeles, 1961.
- Fishbein, M. An investigation of the relationships between beliefs about an object and the attitude toward that object. Human Relations, 1963, 16, 233-240.
- Fishbein, M. A behavior theory approach to the relations between beliefs about an object and the attitude towards the object. In M. Fishbein (Ed.), Readings in attitude theory and measurement, New York: Wiley, 1967, 389-400.
- Fishbein, M., & Hunter, R. Summation versus balance in attitude organization and change. Journal of Abnormal and Social Psychology, 1964, 69, 505-510.
- Fishbein, M., & Raven, B. H. The AB scales: An operational definition of beliefs and attitude. Human Relations, 1962, 15, 35-44.
- Ginter, J. L. An experimental investigation of attitude change and choice of a new brand. Journal of Marketing Research, 1974, 11, 30-40.
- Greene, J. D., & Stock, J. S. Brand attitudes as measures of advertising effects. Journal of Advertising Research, 1966, 6(2), 14-22.
- Hansen, F. Consumer choice behavior: An experimental approach. Journal of Marketing Research, 1969, 6, 436-443.
- Hovland, C. I., Janis, I. L., & Kelley, H. H. Communication and persuasion. New Haven, Conn.: Yale University Press, 1953.
- Howard, J. A., & Sheth, J. N. The theory of buyer behavior. New York: Wiley, 1969.
- Kiesler, C. A. Choice, reward and attitude change. An unpublished paper, 1965.
- Kiesler, C. A., Collins, B. E., & Miller, N. Attitude change: A critical analysis of theoretical approaches. New York: Wiley, 1969.
- Lavidge, R. J., & Steiner, G. A. A model for predicting measurement of advertising effectiveness. Journal of Marketing, 1961, 25 (4), 59-62.
- Maltzman, I. Theoretical conceptions of semantic conditioning and generalization. In T. Dixon & D. Horton (Eds.), Verbal behavior and general behavior theory. Englewood Cliffs, N. J.: Prentice-Hall, 1968.
- Mitchell, A. A. The use of expectancy models for market segmentation. Working series in marketing research, College of Business Administration, Pennsylvania State University, 1974, Paper No. 25.
- Myers, J. H., & Alpert, M. I. Determinant buying attitudes: Meaning and measurement. Journal of Marketing, 1968, 32 (4), 13-20.
- Newell, A., & Simon, H. A. Human problem solving. Englewood Cliffs, N. J.: Prentice-Hall, 1972.

- Nicosia, F. M. Consumer decision processes. Englewood Cliffs, N. J.: Prentice-Hall, 1966.
- Olson, J. C. Cue properties of price: Literature review and theoretical considerations. Working series in marketing research, College of Business Administration, Pennsylvania State University, 1974, Paper No. 20.
- Orne, M. T. On the social psychology of the psychological experiment: With characteristics and their implications. American Psychologist, 1962, 17, 776-783.
- Osgood, C. E. The nature and measurement of meaning. Psychological Bulletin, 1952, 49, 197-237.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. The measurement of meaning. Urbana, Ill.: University of Illinois Press, 1957.
- Osgood, C. E., & Tannenbaum, P. H. The principle of congruity in the prediction of attitude change. Psychological Review, 1955, 62, 42-55.
- Peak, H. Attitude and motivation. In M. Jones (Ed.), Nebraska symposium on motivation. Lincoln: University of Nebraska Press, 1955.
- Platt, J. R. Strong inference. Science, 1964, 146, 347-353.
- Rhine, R. J. A concept-formation approach to attitude acquisition. Psychological Review, 1958, 65, 362-370.
- Rosenberg, M. J. Cognitive structure and attitudinal affect. Journal of Abnormal and Social Psychology, 1956, 53, 367-372.
- Scott, W. A. Attitude change through reward of verbal behavior. Journal of Abnormal and Social Psychology, 1957, 55, 72-75.
- Scott, W. A. Attitude change by response reinforcement: Replication and extension. Sociometry, 1959, 22, 328-335.
- Shapiro, D., & Tagiuri, R. Some effects of response context on trait inferences. Journal of Personality, 1958, 26, 42-50.
- Sheth, J. N. Reply to comments on the nature and uses of expectancy-value models in consumer attitude research. Journal of Marketing Research, 1972, 10, 462-465.
- Sheth, J. N., & Park, C. W. A theory of multidimensional brand loyalty. In S. Ward & P. Wright (Eds.), Advances in consumer research: Volume I. Urbana, Ill.: Association for Consumer Research, 1973, 449-459.
- Staats, A. W. An outline of an integrated learning theory of attitude formation and function. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: Wiley, 1967, 373-376.
- Staats, A. W., & Staats, C. K. Effect of number of trials on the language conditioning of meaning. Journal of General Psychology, 1959, 61, 211-223.
- Staats, A. W., Staats, C. K., & Biggs, D. A. Meaning of verbal stimuli changed by conditioning. Journal of General Psychology, 1958, 71, 429-431.

- Staats, A. W., Staats, C. K., & Crawford, H. L. First-order conditioning of meaning and parallel conditioning of a GSR. Journal of General Psychology, 1962, 67, 159-167.
- Staats, C. K., & Staats, A. W. Meaning established by classical conditioning. Journal of Experimental Psychology, 1957, 74-80.
- Tagiuri, R. Person perception. In G. Lindzey and E. Aronson (Ed.), The handbook of social psychology. Vol. III. Reading, Mass.: Addison-Wesley, 1968.
- Talarzyk, W. W. A reply to the response to Bass, Talarzyk, and Sheth. Journal of Marketing Research, 1972, 9, 465-467.
- Thurstone, L. L. Attitudes can be measured. Journal of Sociology, 1928, 33, 529-554.
- Triandis, H. C. Attitude and attitude change. New York: Wiley, 1971
- Vroom, V. H. Work and motivation. New York: Wiley, 1964.
- Wilkie, W. L., & Pessemier, E. A. Issues in marketing's use of multi-attribute attitude models. Journal of Marketing Research, 1973, 10, 428-441.
- Winter, F. W. A laboratory experiment of individual attitude response to advertising exposure. Journal of Marketing Research, 1973, 10, 130-140.
- Wright, P. L. The cognitive processes mediating acceptance of advertising. Journal of Marketing Research, 1973, 10, 53-62.
- Zanna, M. P., Kiesler, C. A., & Pilkonis, P. A. Positive and negative attitudinal affect established by classical conditioning. Journal of Personality and Social Psychology, 1970, 14, 321-328.

THE EXTENDED FISHBEIN MODEL: ADDITIONAL  
INSIGHTS...AND PROBLEMS

Michael J. Ryan                      E. H. Bonfield  
The University of Alabama

The model was derived from Dulany's theory of propositional control as a means of studying relationships among attitude, behavior, and other variables. Behavioral intention is viewed as a mediator between behavior and attitude and social influence. Theoretical development and research testing the model are reviewed. In addition to showing support for the model, methodological shortcomings, in particular selection of attitude and social influence measures and use of single item measures, are described. The measures which appear to furnish the best statistical fit may not provide the greatest diagnostic power. A reconceptualization of the model indicating causal links is proposed providing a possibility for both best fit and high diagnostic power.

#### Introduction

Much of the discussion and issues involved in the controversy surrounding multi-attribute attitude models has resulted from interest in Rosenberg's (1956) attitude model which has been rekindled largely through the work of Fishbein (1963). Although a great deal of attention has been given the attitude model, as attested by the review of Wilkie and Pessemier (1973), relatively little has been published relative to Fishbein's (1967) extended or intentions model which is an adaptation of Dulany's (1968) theory of propositional control.

Interest in Fishbein's extended model is twofold. First, the model provides a basis for studies of the relationships between attitudes and social influence variables relative to behavior. Second, the model may be useful for the prediction of behavior utilizing behavioral intentions as a mediator. Much of the early research based on the intentions model was encouraging since intentions and behavior were frequently correlated at levels above .85 (Fishbein, 1973).

The purpose of this paper is to review the theoretical framework of the extended Fishbein model along with the research concerning it. The primary focus is on the issues associated with use of the model in consumer research, in particular the conceptual antecedents which underlie the model.

#### Theoretical Development

The Fishbein model is an adaptation of the central theoretical statement contained in Dulany's theory of propositional control. Fishbein's adaptation basically extended the laboratory work of Dulany to a social psychological framework. In order to identify the model's conceptual antecedents Dulany's work is briefly outlined and Fishbein's adaptation explained. The operational procedures used in marketing and social psychology are then discussed.

### Dulany's Theory of Propositional Control

Dulany's theory forms a network of propositions about the effect of reinforcement parameters on verbal responses. Knowledge, beliefs, and awareness are represented as propositions in the form of two hypotheses: (1) the Response Hypothesis (RH), that is, the individual's hypothesis concerning the expectation of a reinforcement, and (2) the Behavioral Hypothesis (BH), that is, the individual's hypothesis concerning the congruence of a response with group norms. Associated with the Response Hypothesis is a feeling of value called the Subjective Value of a Reinforcer (RSv). Likewise, associated with the Behavioral Hypothesis is an evaluative feeling termed the Motivation to Comply (MC). Dulany contends there are many additional variables which affect behavior. However, these other variables have only an indirect effect. They are exogenous to his model and their influence on behavior is reflected in the model's endogenous variables.

The variables included by Dulany reflect specific actions and situations and are proposed to predict and explain behavioral intention (BI). To the extent that the independent variables are specific to a given act, behavioral intention approximates overt behavior.

Through a lengthy process of deductive reasoning, Dulany models his theory to predict an individual's particular verbal response or class of responses in a given situation as

$$B \approx BI = [(RHd) (RSv)] w_0 + [(BH) (MC)] w_1$$

where B = overt behavior; BI = behavioral intention; RHd = hypothesis of the distribution of reinforcement, i.e., the degree to which the individual thinks a specific response will lead to reinforcement or reward; RSv = the subjective value of a reinforcer, i.e., the value the individual places on the reward; BH = behavioral hypothesis, i.e., the degree to which the individual believes a particular behavior is expected of the individual by some specified or generalized set of others; MC = motivation to comply, i.e., the degree of the individual's desire to conform to a BH; and  $w_0$  and  $w_1$  are empirically determined weights.

Dulany's empirical work was carried out in laboratory experiments concerned with verbal conditioning and concept attainment. As such, there was only one RHd and BH and their respective RSv and MC. These variables were generally manipulated by an experimenter who, within the context of the experiment, was the only influence source. Dulany (1968) reports several tests of the model. These tests support the assumption of additivity of independent variables as well as the need to include BI as a moderator. The independent variables accounted for a large proportion of the variance in BI (50-77%) and BI accounted for a large proportion of the variance in B (80-88%).

### The Fishbein Adaptation

Fishbein restated Dulany's theory as follows:

An individual's intention to perform a specific act, with respect to a given stimulus object, in a given situation, is a function of the following:

- (1a.) His beliefs,  $B_i$ , about the consequences of performing a particular behavior (in a given

situation), that is, the probability or improbability that the performance of behavior  $x$  will lead to some consequence  $y_i$ .

- (1b.) The evaluative aspect,  $a_i$ , of  $B_i$ , that is, the S's evaluation of  $y_i$ .
- (2a.) A normative belief, that is, the S's belief about what he should do in this situation (NB).
- (2b.) His motivation to comply with the norm, that is, his desire, or lack of desire, to do what he thinks he should do (MC). (Fishbein, 1967, p. 488)

In the above restatement, item (1a) or  $B_i$  approximates Rhd, item (1b) or  $a_i$  approximates Rsv, item (2a) or NB approximates BH, and item (2b) approximates MC. Fishbein thus shows the first component, (Rhd) (Rsv), of the Dulany theory to be analogous to the summative attitude model associated with him

$$A_{act} = \sum_{i=1}^n B_i a_i$$

where  $A_{act}$  = the attitude toward performance of a specific act, and  $B_i$  and  $a_i$  have been previously defined in Fishbein's restatement of Dulany's theory. The Aact model differs from Fishbein's  $A_o$ , or attitude toward an object model, more popularly described in the marketing literature. In the  $A_o$  model,  $B_i$  statements refer to concept objects which, in marketing, are usually product attributes.  $B_i$  in the Aact model, however, is associated with behavioral outcomes. Fishbein (1971) has stated the Aact conceptualization is more appropriate for predicting and understanding purchase intentions and behavior. The argument is analogous to the features-benefit position traditionally taken in marketing (Haley, 1968).

Fishbein's adaptation of Dulany's second component, (BH) (MC), is more general. Social normative beliefs (SNB) are described as

...the individual's beliefs about what "society" (i.e., most other people, his "significant others," etc.) "says" he should do (i.e., a social or group norm). (Fishbein, 1967, p. 489)

This conceptualization of the second component is an expansion of Dulany's work since he conceived this variable in terms of pressure exerted by the experimenter.

Fishbein summarizes the entire adaptation

To summarize briefly then, it can be seen that in its adapted form, the theory essentially leads to the prediction that an individual's intention to perform any behavior in a given situation (and thus his actual performance of the behavior) is a function of (1) his attitude toward performing the behavior in the situation, (2) his perception of the norms governing that behavior in that situation, and (3) his motivation to comply with



those norms. (Fishbein, 1972a, p. 248)

This formulation is modeled

$$B \approx BI = [Aact]w_0 + [NB \cdot MC]w_1 \quad (1)$$

where NB = a normative belief, the degree of belief that others expect the individual to perform a specific act, MC = motivation to comply with the expectations of others. All other elements have been previously defined.

Although Fishbein does not use the term, BI appears to be a moderating variable. It comes between the independent and dependent variables in a causal sequence. It is viewed as a consequence of the independent variables and as a determinant of the dependent variable.

The model purports to predict and explain human behavior. Fishbein (1973) states that the model is appropriate for a number of behavioral criteria. Applications of the model are very restrictive in its present form due to a set of exogenous variables.

Generally speaking, there are three major factors that influence the size of the relationship between intention and behavior: (1) the specificity of the intentional measure; (2) the time between the measure of intention and the behavioral observation; and (3) the degree to which carrying out the intention is completely under the individual's control. (Fishbein, 1973, p. 15).

Fishbein considers the first of these factors to be of primary importance. To put it more simply, the more removed the measure of intention is from the criteria to be predicted, the poorer will be the prediction. The second and third restrictions have also been considered in the more general consumer behavior models (e.g., Howard and Sheth, 1969; Andreason, 1965) and by surveyors of consumer intentions to purchase durables (Clawson, 1971; Juster, 1966). To date, no attempts have been made to make these variables endogenous to Fishbein's model. Also, the attitudinal and social influences must be situation specific. It is not expected that general attitude or social influence measures will predict behavioral intention.

In addition to form (1), Fishbein has offered two basic alternative formulations of the model. The first alternative form involves a departure from Dulany's conceptualization in which the possibility of adding a third component, personal normative beliefs (NBp), is considered. This component is described as

The individual belief about what he personally feels he should do (i.e., a personal norm or rule of behavior).  
(Fishbein, 1967, p. 489)

This modified formulation is

$$B \approx [Aact]w_0 + [NBs \cdot MCs]w_1 + [NBp \cdot MCP]w_2 \quad (2)$$

where p and s refer to personal and social norms respectively. Fishbein's addition of the personal norm component was his initial method of handling the

question of whether NB referred to a personal norm, a social norm, or both.

The second alternative form of the extended Fishbein model results from recognition that NBs may result from several sources.

This formulation suggests that it may be necessary to consider many different types of normative beliefs; for example; beliefs about what one's (a) parents, (b) friends, (c) co-workers, (d) religious group, etc., "says" the individual should do. (Fishbein, 1967, p. 490).

This formulation is modeled

$$B \approx BI = [Aact]w_0 + \left[ \sum_{j=1}^k NB_j MC_j \right] w_1 \quad (3)$$

where  $NB_j$  = degree of belief that a specific act is expected by the  $j$ th group or person and  $MC_j$  = motivation to comply with the expectation of the  $j$ th group or person.

Model forms (1), (2), and (3) represent the basic extended Fishbein model and the original alternative forms he suggested. Operationalization procedures have lead to additional forms of the model.

The forms of the model discussed state the independent variables combine additively. That is, they are orthogonal. Conceptually, these constructs may have separable effects on BI. On the other hand, these constructs may not be completely independent. The notion of perceptual distortion, for example, is based on an interrelationship between attitude and social influences.

#### Operationalization Procedures

Fishbein's operationalization procedures follow two basic steps: (1) determining the salient  $B_i$  outcomes and  $NB_j$  groups and (2) developing techniques for measuring the constructs.

#### Determining Saliency

Fishbein (1963) determines salient behavioral outcomes according to procedures developed by Maltzman, Bogartz, and Breger (1958). Subjects are asked to give different associations to the same stimulus words in a free association situation. When Aact measures are being developed, the stimulus words refer to behaviors rather than attitude objects. The technique precludes probing, prompting, or presentation of predetermined lists of behavioral outcomes or potential associations in order to avoid triggering sets of interdependent beliefs based on social desirability rather than true feelings (Cowling, 1973b; Fishbein, 1971). Based on research on information processing and span of attention, Fishbein (1971) states there are probably only 5 to 9 behavioral outcomes that serve as the primary determinants of a person's attitude at any point in time.

...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.

...If we really want to know the determinants of attitude we have to know the person's salient beliefs. Unfortunately, there is at the present time no independent way of assessing salience outside of order of elicitation, and even more importantly, a non-salient belief may be just as good as an indicant of a person's attitude as a salient belief. Thus at present, outside of using a direct elicitation procedure, there is no way of telling whether you've obtained salient or non-salient beliefs. (Fishbein, 1971, pp. 313-314).

The relevant beliefs, or behavioral outcomes, and reference groups, when determined according to the elicitation procedure, are obtained from the same population being studied (e.g., Jaccard and Davidson, 1972). Fishbein believes this procedure is more theoretically sound than either "in depth probing" or using factor or regression analysis to analyze a predetermined list of attributes or outcomes.

Construct Operationalization

Behavior (B), has been generally operationalized as the observation of an individual's choice in a specific situation. For example, Bonfield (1974) used consumer dairies to obtain self reports of fruit drink purchases. Harrell and Bennett (1974) used prescription records provided by a trade association.

The procedures used to measure the remaining constructs have been based on semantic differential techniques (Osgood, Suci, and Tannenbaum, 1957).

Behavioral intention (BI) was usually operationalized with a seven place scale from probable to improbable. For example (Jaccard and Davidson, 1972)

I intend to use birth control pills

probable \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ improbable

A disadvantage here is that such one item instruments tend to be unreliable. Lutz (1973b) used multiple item scales such as

When it is introduced onto the market, I intend to try Brand M detergent.

likely \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ unlikely  
 probable \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ improbable  
 possible \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ : \_\_\_\_\_ impossible

When multiple item scales are utilized, it is possible to sum scores over all scales or take an average as either scoring method will yield the same correlations. Averages are preferable if results are to be compared with studies using single item scales.

Aact has been operationalized as the sum over four semantic differential type items. For example (Ajzen and Fishbein, 1970)



$\Sigma B_i a_i$  and Aact scores.

The majority of these studies operationalized the normative component as the sum of separate sources of influence, see (3), with two studies (Ajzen and Fishbein, 1969 and 1970) utilizing generalized other as a single normative referent.

The research reviewed by Ajzen and Fishbein does not clearly indicate whether NBp should be included in the model. Ajzen and Fishbein (1973) indicated empirical findings have shown NBp to be an alternative measure of BI. In recent writings Fishbein (e.g. 1973) only described the model in form (3) which excludes NBp. Also, Ajzen and Fishbein (1969) reported significant beta weights for Aact, NBp and NBs supporting the model with three independent variables (2). However, in a marketing application, Bonfield (1974) found NBp to act as a suppressor variable that caused Aact to become statistically insignificant.

Although model (3) appears to be the appropriate model for adaptation in marketing, the operationalization of Aact as a four-item semantic differential type scale rather than as  $\Sigma B_i a_i$  reduces the diagnostic power of the model.

#### Studies in Marketing

A number of marketing studies testing the model have been conducted in Great Britain (Tuck, 1973). The majority of the work has been done commercially, however, and detailed results have been held as proprietary. Summary results from five of these studies have been published (Cowling, 1973a, b; Bright and Stammers, 1971; Bruce, 1971; and Tuck and Nelson, 1969). The results of these studies suggest elicitation of salient behavioral outcomes for specific segments and products results in higher correlations between Aact and  $\Sigma B_i a_i$  than using predetermined lists. Aact was more highly correlated with BI than either  $A_0$  or  $\Sigma B_i a_i$ . These findings support Fishbein's contention that Aact is a more appropriate predictor of BI than  $A_0$  and that elicitation techniques are necessary. Finally, the British studies showed relative weights associated with attitudinal and normative components of the model vary across products and market segments. Stronger correlations between BI and Aact versus BI and  $A_0$  have also been reported by American market researchers (Weddle and Bettman, 1973; Mathews, et. al., 1974).

Nine U.S. studies have been reported in which a form or adaptation of the extended model has been tested in a marketing context. The major results of these studies are summarized in Table I. The model appears to have value in predicting and explaining variance in intentions and behaviors over a wide range of purchase intentions and purchase behavior. Harrell and Bennett (1974) tested the model with respect to physician prescribing behavior, Ryan (1973) with respect to intentions toward purchase of products of automobile manufacturers, Ryan (1974), Bonfield (1974), Lutz (1973), and Mathews, Wilson and Harvey (1972) with respect to consumer convenience goods, Weddle and Bettman (1973) with respect to purchasing underground term papers, and Wilson, Mathews and Monoky (1973) with respect to bargaining in a personal selling situation. Harrell and Bennett (1974), Bonfield (1974), Mathews, et. al. (1974), and Wilson, et. al. (1973) reported correlations between intentions and behavior while others did not report measuring behavior.

The average correlation between BI and B across these studies was .435.

Table 1  
Summary of Findings from Initial Marketing Applications  
of Fishbein's Behavioral Intention Model

Study	N	Beta Weights		Multiple Correlations on BI	Correlation BI, B
		Aact	$NB_jMC_j$		
Harrell and Bennett (1974) <sup>a</sup>					
Prescription Brand A	93	.337	.175	.424**	.517**
Prescription Brand B	93	.393	.076	.409**	.472**
Prescription Brand C	93	.332	.228	.450**	.274**
Prescription Brand D	93	.387	.151	.461**	.270**
Prescription Brand E	93	.467	.202	.539**	.336**
Mathews, et. al. (1974)					
Gleem	162	.249**	.463**	.620**	
Crest	162	.219**	.546**	.668**	
Colgate	162	.363**	.386**	.673**	
Ultra Brite	162	.111*	.592**	.664**	
MacLeans	162	.127	.578**	.655**	
Pepsodent	162	.099*	.683**	.746**	
Bonfield (1974) <sup>a,b</sup>					
Fruit Drinks					
Total Sample	158	.212**	.444**	.601**	.402**
High Importance	52	.409**	.475**	.811**	.441**
Low Importance	106	.120	.427**	.499**	.382**
High Loyalty	53	.001	.642**	.623**	.038
Medium Loyalty	34	.248	.195	.332**	.219
Low Loyalty	27	.923**	-.157**	.775**	.415*
Low Education	85	.141	.403**	.493**	.370**
High Education	73	.290**	.493**	.705**	.437**
Ryan (1974) <sup>c</sup>					
Ultra Brite	80	.355**	.203	.475**	
Mustang II	80	.731**	.060	.761**	
Weddle & Bettman (1973) <sup>d</sup>					
Purchasing Term Papers	57	.49 **	.24 *	.54 **	
Ryan (1973)					
Chrysler Corporation	91	.455**	.353**	.676**	
General Motors Corp.	91	.593**	.276**	.734**	
Ford Motor Co.	91	.641**	.169*	.696**	
American Motors	91	.375**	.410**	.648**	
Foreign Mfgs.	91	.576**	.193*	.663**	
Lutz (1973b)					
Fictitious Detergent	246	.646**	.048	.647**	
Fictitious Detergent <sup>d</sup>	246	.465**	.130	.480**	
Lutz (1973a) <sup>d</sup>					
Season Football Tickets	100	.683**	.084	.73 **	
Season Football Tickets	77	.603**	.203**	.75 **	
Wilson, et. al. (1973)					
Prisoners' Dilemma Game					
Perceived Similar	20	.613**	.299	.708**	.823**
Perceived Dissimilar	20	.744**	-.032	.690**	.660**

\*p &lt; .05

\*\*p &lt; .01

<sup>a</sup>Aact was operationalized using a factor analytic scheme suggested by Howard and Sheth (1969).<sup>b</sup>Social normative influence was operationalized as a 'general other'.<sup>c</sup>Social normative influence was operationalized as Aact<sub>0</sub>.<sup>d</sup>Aact was operationalized as  $\sum_{i=1}^n B_i a_i$

The average multiple correlation of attitude and social influence on BI was .62. Three of these studies (Harrell and Bennett, 1974; Mathews, et. al., 1974; Lutz 1973a) also employed cross validation procedures. Correlations obtained using cross validation were comparable to those above indicating the model is a stable predictor.

Using  $R^2$  as a criterion, the predictive power obtained in the marketing studies has been generally lower than obtained in the social psychology studies. A possible explanation may be found in the different types of behaviors and attitudes examined. The social psychology studies were predominantly concerned with potentially central attitudes (e.g., attitudes toward behavior involving racial or religious beliefs). The marketing studies examined purchase activities which may have involved noncentral attitudes. Consequently, the marketing behavior may have had smaller associations with attitudes and social influence because of their lack of centrality to the individual. Bonfield (1974) found the deterministic influence was stronger when subjects perceived the product as important. Ryan (1974) also found the deterministic influence was stronger for automobiles than toothpaste. An automobile purchase would be expected to be more central than a toothpaste purchase.

In an attempt to explain why marketing behavioral models have generally accounted for a small proportion of systematic variance, Bass (1974) has demonstrated that purchase behavior is predominantly a stochastic process. The emerging pattern in the studies discussed here suggests the relative influence of deterministic and stochastic elements may vary across products.

Bonfield (1974) also found the relative influence varied across different segments. In addition to perceived product importance, high education and low brand loyalty groups exhibited stronger deterministic influences than low education and high brand loyalty groups respectively. Wilson, et. al. (1974) accounted for more variation in housewives' purchase intentions toward toothpaste brands than did Ryan (1974) among college undergraduates. The housewives may have more carefully considered toothpaste purchase than college students.

These findings suggest claims that behavior is primarily predicted from either deterministic or stochastic influences may be premature. Rather, the relative importance of these influences may be a function of group, individual, or product characteristics. In general, the  $R^2$ 's found in the marketing studies have not shown the model to be a particularly good predictor. Bonfield (1974) found a naive model predicted equally well. Instead, the patterns in relative  $R^2$ 's associated with the model suggest the model has explanatory potential. A synthesis of a much larger body of research is needed to suggest a more definitive pattern of deterministic and stochastic influences among products and segments. The strategic implications are obvious. For example, promotional strategies based on attitude change tactics are inappropriate for products and groups whose purchase behavior is primarily stochastic.

### Operationalization Differences

Operationalization of the dependent and independent components of the model has varied considerably in the marketing studies conducted in the United States. Behavior has been measured in field studies in terms of the actual brand most frequently prescribed by physicians over a period running both before and after measures of intention, attitude, and social influence had been obtained (Harrell and Bennett, 1974), brand of fruit drink purchased on the first occasion following the independent component measures (Bonfield, 1974), and choice of a

free sample immediately following an interview (Mathews, et. al., 1974). Only Wilson, et. al. (1973) measured behavior as more than a dichotomous outcome, treating the variable in terms of cooperative choices in a prisoner dilemma, negotiation experiment.

Attitude has been measured closely paralleling Fishbein's Aact procedures as well as in terms of substituting other measures. As noted previously, Fishbein has used  $A_0$ ,  $\Sigma B_i a_i$ , and Aact (evaluative semantic differential type scales) in various social psychological studies. Aact was operationalized as a multi-item, semantic differential type scale by Mathews, et. al. (1974), Ryan (1973), and Wilson, et. al. (1973); Harrell and Bennett (1974), and Weddle and Bettman (1973) basically operationalized attitudes as  $\Sigma B_i a_i$ . Ryan (1974) and Lutz (1973b) used both techniques. Bonfield (1974) substituted a factor analytic model as described by Howard Sheth (1969).

A  $\Sigma B_i a_i$  operationalization of the attitude component appears essential to understanding since it represents the structure underlying attitude. However, operationalization in this form has presented some difficulties. Among the marketing studies using  $\Sigma B_i a_i$  scales only Ryan (1974) used multi-item  $B_i$  and  $a_i$  scales. One item scales are generally considered unreliable since there is no opportunity for random and specificity errors to average out over a set of scales. Reliability estimates should be incorporated in future correlational studies in order to ascertain whether weak correlations are due to unreliable measures.

Semantic differential type measurements deviate from the interval and ratio characteristics necessary for their applications in expectancy models (Heise, 1969). Although violations of the interval assumptions make little difference when correlating scale scores (Nunnally, 1967, pp. 24-30), multiplication of these scores does require a ratio scale in order to be meaningful (Lord and Novick 1968, p. 21). Schmidt (1973) addressed the multiplicity issue and provided empirical evidence indicating expectancy value models are not robust with respect to violating ratio scale assumptions. Scaling and operational procedures are an important area for further research which consumer researchers are beginning to explore (e.g. Bettman, Capon, and Lutz, 1974). Future studies should incorporate both Aact and  $\Sigma B_i a_i$  operationalizations in order to provide both appropriate correlation coefficients, beta weights, and understanding the underlying attitudinal structure.

While there is general agreement the first component of the extended model is attitude, some confusion exists as to the theoretical conceptualization of the second component. Fishbein (1967) originally conceived of it as a social normative component. More recently he has viewed it as a measure of the perceived attitude of others toward performing the behavior (Ajzen and Fishbein, 1972) as has Ryan (1974) in a marketing context. Bonfield (1974) referred to the component as an indicant of reference group or social influence.

Four strategies have been followed in operationalizing the social influence component of the model. Wilson, et. al. (1973), following an experimental strategy paralleling Dulany's work, controlled the situation so there would be only one  $NB_j MC_j$  source. Bonfield (1974) and Harrell and Bennett (1974) used single item, generalized other  $NB_j MC_j$  operationalizations. Mathews, et. al. (1974); Ryan (1973, 1974), Lutz (1973), and Weddle and Bettman (1973) utilized multi-other,  $NB_j MC_j$  scales where each  $NB \cdot MC$  represented a specific source of normative influence. Ryan (1974) elicited salient others as influence sources. The other studies appeared to use predetermined lists.



Ryan (1974), following Ajzen and Fishbein (1972), also operationalized the social influence component as perceived attitude of relevant others,  $Aact_0$ , toward performing the behavior. Ajzen and Fishbein (1972) have conceived of  $Aact_0$  as the attitude of others toward their, the others, performing the act, not their attitude toward the individual, whose intentions are the subject of the model, performing the act. Essentially, the Ajzen and Fishbein statement is of the form

Most of the people whose opinion is important to me with  
respect to this act think this act is

while Ryan (1974) has operationalized the statement in the form

Others think that for me to perform this act is

The same evaluative semantic differential type scales, good-bad, foolish-wise harmful-beneficial, and punishing-rewarding, have been used by both Ryan and Ajzen and Fishbein.

There is no apparent behavioral expectation element in the  $Aact_0$  component as operationalized by Ajzen and Fishbein, although the expectation notion appears essential in the antecedent, Dulany theory. In addition, their  $Aact_0$  measure is inconsistent with Fishbein's conceptualization of Aact which requires the attitude to be linked to the individual's performance of some behavioral act. Ryan's operationalization is consistent with the expectancy notion in the Dulany theory and the specificity requirement of the Fishbein Aact conceptualization.

Disturbingly, operationalizing the second component as  $Aact_0$  and as  $NB_jMC_j$  may not be equivalent. Ryan (1974) has shown correlations of less than .05 between the two measures.  $Aact_0$  correlated more highly with BI than did  $NB_jMC_j$  or than  $Aact_0$  correlated with  $\Sigma NB_jMC_j$ . Thus,  $Aact_0$ , operationalized as other's attitude toward the perceiver's performance may be an appropriate operationalization of the second component when prediction is the purpose of the model. The  $NB_jMC_j$  method of operationalization, however, provides better diagnostic power.

Additional study of the second component of the model is needed. Additional evidence is needed relative to the inclusion of MC, the substitutability of  $Aact_0$  and  $NB_jMC_j$ , and construct validity of these components. Since Aact and  $B_ja_j$  measurement problems also apply here, it is necessary to have the same type of research scrutiny paid the normative or social influence measures.

#### Additivity of Independent Variables

The forms of the model discussed state the independent variables combine in an additive manner. That is, they are orthogonal. Operational measure of these constructs have been shown to have separate effects on BI. Yet, common

sense suggests these constructs are not completely independent. For example, a person who perceives an act as morally correct would be expected to believe others also view this act as morally correct. Thus, on a theoretical basis it is expected that interaction as well as direct effects should be present in the model.

The number of statistically significant beta weights for  $A_{act}$  and  $ENB_j MC_j$ , shown in Table 1 far exceed the number that are insignificant, supporting the contention that each independent variable has direct predictive power. This evidence must be viewed with caution since only Ryan (1974) used measures for both attitude and social influence that were highly and equally reliable. In Ryan's research, the beta weights associated with social influence were not statistically significant.

Evidence suggesting the independent components are not additive is shown by high correlations between independent variables. Bonfield (1974), Mathews, *et. al.* (1974), and Ryan (1974) found correlations between attitude and social influence which were higher than the correlation of either on BI. In an experimental setting, Ryan (1974) found changes in both  $A_{act}$  and  $A_{act_0}$  occurred regardless of whether  $B_i$  or  $NB_j$  were manipulated, thus suggesting a nonadditive relationship. More research is needed to investigate the additivity assumption. In addition nonadditive reformulations of the model based on intuitively appealing conceptual antecedents should be investigated.

If an additive model can be supported, beta weight analysis has implications for marketing strategies as a means of ascertaining whether brand or product purchase intentions are primarily under attitudinal or social influence control. For example, Mathews, *et. al.* (1974) suggested an attitudinal-social influence continuum exists analogous to an instrumental-expressive continuum used for classifying consumer products. That is, automobiles and clothing were classified as expressive products since they are indicants of social status. Thus, automobiles and clothing would be expected to have higher beta weights associated with social influence than with attitude. Laundry detergents were viewed as instrumental products in that they are relatively homogeneous, generally purchased and used for utilitarian benefits, and do not tend to carry status implications. Therefore, beta weights associated with attitude toward instrumental products would be expected to be greater than the beta weights associated with social influence. Mathews, *et. al.* caution that the instrumental and expressive dimensions are not mutually exclusive. The results reported by Mathews, *et. al.* (1974) support their hypothesis. Variation in intentions concerning toothpaste purchases explained by attitude were consistently lower than explained by social influence, but the beta weights associated with attitude for the cosmetic brands (Ultra Brite, MacLeans, and Pepsodent) were lower than the beta weights associated with attitude for the noncosmetic brands and social influence was relatively stronger with respect to the cosmetic brands (See Table 1). Wilson, *et. al.* (1973) found situational influences affected the relative importance of attitudes and social influences. Attitudinal influences predominated, but social influences were stronger among perceived similar buyer-seller dyads.

#### The Relationship Between Behavior and Behavioral Intentions

Sheth (1974) has defined behavior as a function of behavioral intention and those situational factors that could not be predicted by the individual at the time of verbally expressing his behavioral intention. In the studies where BI and B measures were contiguous or nearly contiguous, BI-B correlations were high (Mathews, *et. al.* 1974; and Wilson, *et. al.* 1973). In the studies where

measures of B and BI were allowed to vary, the correlations were much lower (Bonfield, 1974; and Harrell and Bennett, 1974). Future studies should examine situational variables intervening between measures of B and BI. Perhaps the procedures developed by Sheth (1973) could be adapted to the present model.

Three social psychology studies (Darroch 1971; Fishbein, *et. al.* 1970; Ajzen and Fishbein, 1970) have furnished evidence indicating the need to include BI as a moderator variable even though situational variables were controlled. These findings suggest understanding purchase intentions may be necessary for understanding purchase behavior. Consumer researchers should address this issue.

### A Reconceptualization of the Model

The theoretical development of the extended Fishbein model along with the empirical evidence gathered testing it leads to a number of causal inferences. Consistent with Fishbein's earlier work, the independent attitude variable is  $A_{act}$  not  $\Sigma B_i a_i$ . Evaluation and beliefs explain, that is, come before  $A_{act}$ . Although Fishbein does not use the term antecedent variable,  $\Sigma B_i a_i$  fits the use prescribed by Rosenberg (1968) for this term, that is, attempting to discover first causes.

The antecedent variable is a true effective influence; it does not explain away the relationship between the independent and dependent variables but clarifies the influences which precede this relationship. (Rosenberg, 1968, p. 66)

Following the same reasoning used to describe the two forms of attitude as antecedent and independent variables, it appears reasonable to suggest  $\Sigma N B_j M C_j$  represents an underlying structure antecedent to  $A_{act}$ .

Although Fishbein does not use the term, BI appears to be a moderating variable. It comes between the independent and dependent variables in a causal sequence. It is viewed as a consequence of the independent variables and as a determinant of the dependent variable. That is, it is a necessary link in the causal chain.

Finally, the empirical evidence reviewed here suggests substantial interdependence among the variables. That is, changes in  $\Sigma N B_j M C_j$  are not only expected to lead to changes in  $A_{act}$  and BI, but also in  $A_{act}$  while changes in  $B_i a_i$  are not only expected to lead to changes in  $A_{act}$  and BI, but also in  $A_{act}$ . In addition, changes in  $A_{act}$  are expected to lead to changes in  $A_{act}$  and vice versa. These causal influences are indicated in Figure 1.

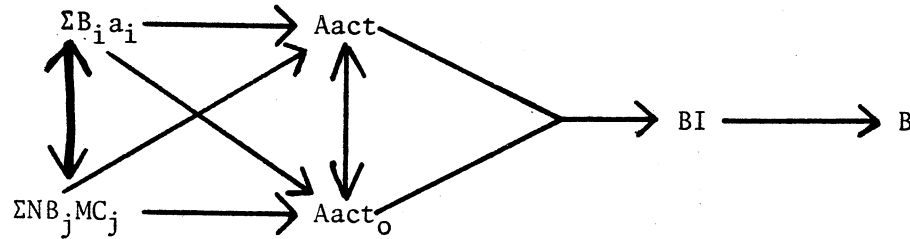
There are three generally accepted demonstrations inferring causal relationships. First, there should be an association between the independent and dependent variables. Second, there should be causal order. That is, the independent variables should be shown to occur prior to the dependent variables. Third, the association should not be spurious. That is, the association between variables should be shown not to disappear when the influences of rival causal variables are removed. One has only to break one of these three links to establish noncausality (Popper, 1959).

In order to test this reconceptualization of the model, two research strategies are suggested. The first involves the use of path analysis as a means of determining whether empirical evidence will support the relationships in

Figure 1. Secondly, experimental designs should be utilized to test the causal order assumptions which underlie path analysis. If the associations and causal order can be supported, competing variables should be introduced as antecedents in order to test for spuriousness in the relationships.

Figure 1

The Suggested Causal Sequence in the Extended Fishbein Model



#### Suggested Future Research

One avenue of potential future research has already been suggested in terms of the reconceptualization of the model. In addition, the methodological shortcomings described throughout the discussion of empirical evidence lead to a call for better research conceptualization where possible.

Specifically, whenever the underlying structure of Aact and Aact<sub>0</sub> are being studied, it seems necessary to elicit behavioral outcomes and salient others for each behavior and homogeneous target group.

One item scales, being potentially unreliable, should be replaced by multi-item scales for B<sub>i</sub>, a<sub>i</sub>, NB<sub>j</sub>, and MC<sub>j</sub>. In addition, care should be taken in use of evaluative, semantic differential type scales to measure Aact and Aact<sub>0</sub>. Only evaluative scales should be used in these measures as ascertained by factor analysis since it is only the evaluative scales that have been shown to correlate with standard attitude measures (Osgood, et. al., 1957). Enough semantic items should be included to adequately represent evaluative semantic space and provide reliable scales.

Validation of attitude and social influence measures are needed. Although, Aact is generally measured using semantic differential type scales, the measure has not been validated utilizing a standard semantic differential bipolar adjective set. In addition, ΣB<sub>i</sub>a<sub>i</sub> measures have never been shown to correlate as high with semantic differential evaluative scales as the latter scales have been shown to correlate with Thrustone equal-appearing interval scales. Unfortunately, validating constructs for the social influence component are not so clearly identifiable since more conceptual work is needed as a starting point.

Few studies have included cross validation techniques in model parameter estimates. Cross validation should be a standard procedure in correlational tests of the extended model in order to check for stability.

Some question still exists relative to the appropriate method for scoring B<sub>i</sub>, a<sub>i</sub>, NB<sub>j</sub> and MC<sub>j</sub> scales. The only work reported testing coding schemes has shown bipolar scales to be superior to unipolar scales while also supporting the multiplication of B<sub>i</sub>a<sub>i</sub> scales (Bettman, Capon, and Lutz, 1974).

The body of research based on the extended Fishbein model is now considerable and the results of tests suggest the model has value for the understanding of a wide range of purchase behavior and behavioral intentions. However, the validation process has just begun. In addition to the possible avenues of research pointed out, replications are needed and the range of purchase behavior and subject and respondent groups needs to be expanded. Hopefully, future research will build on what has been done in the past so a cohesive body of knowledge will continue to develop.

## REFERENCES

- Ajzen, I. & Fishbein, M. The prediction of behavioral intentions in a choice situation. Journal of Experimental Social Psychology, 1969, 5, 400-416.
- Ajzen, I. & Fishbein, M. The prediction of behavior from attitudinal and normative variables. Journal of Experimental Psychology, 1970, 6, 466-487.
- Ajzen, I. & Fishbein, M. Attitudes and normative beliefs as factors influencing behavioral intentions. Journal of Personality and Social Psychology, 1972, 21, 1-9.
- Ajzen, I. & Fishbein, M. Attitudinal and normative variables as predictors of specific behaviors. Journal of Personality and Social Psychology, 1973, 27, 41-57.
- Andreason, A. R. Attitudes and consumer behavior: A decision model. In L. Preston (Ed.), New research in marketing. Berkeley: University of California, Institute for Business and Economic Research, 1965, 1-16.
- Bass, F. M. The theory of stochastic preference and brand switching. Journal of Marketing Research, 1974, 11, 1-20.
- Bettman, J. R., Capon, N. & Lutz, R. J. Cognitive algebra in multi-attribute attitude models. Working paper no. 10, Center for Marketing Studies, University of California, Los Angeles, February 1974.
- Bonfield, E. H. Attitude, social influence, personal norms, and intention interactions as related to brand purchase behavior. Journal of Marketing Research, 1974, in press.
- Bright, B. E. & Stammers, H. R. Advertising, promotion and medical detailing. Proceedings: Seminar on pharmaceutical research, European Society of Market Research, 1971.
- Bruce, J. First experiences with Fishbein theory and survey methods. Proceedings: Seminar on translating advanced advertising theories into research reality. Amsterdam, The Netherlands: European Society of Market Research, 1971, 25-39.
- Clawson, C. J. How useful are 90-day purchase probabilities? Journal of Marketing, 1971, 35(4), 43-47.
- Cowling, A. B. Determining and influencing consumer purchase decision. European Research, 1973a, 1, 26-31, 34.
- Cowling, A. B. Use of elicitation technique for producing dimension of brand

- choice. Proceedings. British market research society, 1973b, 139-156.
- Darroch, R. K. The prediction of intention and behavior: A comparison of theoretical models of the attitude-behavior relationship using a photographic release signing situation. Unpublished doctoral dissertation, University of Illinois, 1971.
- Dulany, D. E. Awareness, rules, and propositional control: A confrontation with S-R behavior theory. In D. Horton and T. Dixon (Eds.), Verbal behavioral and general behavior theory. New York: Prentice-Hall, 1967, 340-387.
- Fishbein, M. A. An investigation of the relationships between beliefs about an object and the attitude toward that object. Human Relations, 1963, 16, 233-239.
- Fishbein, M. A. Attitude and the prediction of behavior. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: Wiley, 1967, 477-492.
- Fishbein, M. A. Some comments on the use of "models" in advertising research. In Proceedings: Seminar on translating advanced advertising theories into research reality. Amsterdam: The Netherlands: European Society of Market Research, 1971.
- Fishbein, M. A. The search for attitudinal-behavior consistency. In J. B. Cohen (Ed.), Behavioral science foundations of consumer behavior. New York: Free Press, 1972, 245-252.
- Fishbein, M. A. The prediction of behavior from attitudinal variables. In C. D. Mortensen and K. K. Sereno (Eds.), Advances in communication research. New York: Harper and Row, 1973, 3-31.
- Haley, R. I. Benefit segmentation: A decision oriented research tool. Journal of Marketing, 1968, 32 (3), 30-35.
- Harrell, G. D. & Bennett, P. D. An evaluation of the expectancy value model of attitude measurement for physician prescribing behavior. Journal of Marketing Research, 1974, 11, 269-278.
- Heise, D. R. Some methodological issues in semantic differential research. Psychological Bulletin, 1969, 72, 406-422.
- Howard, J. A. & Sheth, J. N. The theory of buyer behavior. New York: Wiley, 1969.
- Jaccard, J. J. & Davidson, A. R. Toward an understanding of family planning behaviors: An initial investigation. Journal of Applied Social Psychology, 1972, 2, 228-235.
- Juster, F. T. Consumer buying intentions and purchase probability: An experiment in survey design. Journal of the American Statistical Association, 1966, 61, 658-696.
- Lord, F. M. & Novick, M. R. Statistical theories of mental test scores. Reading, Massachusetts: Addison Wesley, 1968.
- Lutz, R. J. A comparison of two alternative models of the attitude-behavior

- relationship. Paper presented at the 81st Annual Convention of the American Psychological Association, Division 23, Montreal, Canada, August 27-31, 1973a.
- Lutz, R. J. Cognitive change and attitude change: A validation study. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, 1973b.
- Maltzman, I., Bogartz, W., & Breger, L. A procedure for increasing word association originality and its transfer effects. Journal of Experimental Psychology, 1958, 56, 392-398.
- Mathews, H. L., Wilson, D. T., & Harvey, J. W. An empirical test of the Fishbein behavioral intention model. Working paper no. 19, College of Business Administration, The Pennsylvania State University, February, 1974.
- Nunnally, J. C. Psychometric theory. New York: McGraw-Hill, 1967.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. The measurement of meaning. Champaign: University of Illinois, Press, 1957.
- Popper, K. R. The logic of scientific discovery. New York: Basic Books, 1959.
- Rosenberg, M. J. Cognitive structure and attitudinal effect. Journal of Abnormal and Social Psychology, 1956, 53, 367-72.
- Rosenberg, M. The logic of survey analysis. New York: Basic Books, 1968.
- Ryan, M. J. Purchase behavior: An explanatory theory, a predictive model. In R. L. King (Ed.), 1973 proceedings. Southern Marketing Association.
- Ryan, M. J. An empirical test of a predictive model and causal chain derived from Fishbein's behavioral intention model and applied to a purchase intention situation. Unpublished doctoral dissertation, University of Kentucky, 1974.
- Schmidt, F. L. Implications of a measurement problem for expectancy theory research. Organizational Behavior and Human Performance, 1973, 10, 243-251.
- Sheth, J. N. A field study of attitude structure and attitude-behavior relationship. Faculty working paper no. 116, College of Commerce and Business Administration, University of Illinois at Urbana-Champaign, 1973.
- Sheth, J. N. An investigation of relationships among evaluative beliefs, affect, behavioral intention, and behavior. In J. U. Farley, J. A. Howard, & L. W. Ring (Eds.), Consumer behavior: Theory and application. Boston: Allyn & Bacon, 1974, 89-114.
- Tuck, M. Fishbein theory and the Bass-Talarzyk problem. Journal of Marketing Research, 1973, 10, 345-348.
- Tuck, M. & Nelson, E. N. The relationship between attitudes and behavior. In ESOMAR congress papers. Amsterdam: The Netherlands: European Society of Market Research, 1969, 257-278.

- Weddle, D. E. & Bettman, J. R., Marketing underground: n investigation of Fishbein's behavioral intention model. In S. Ward and P. Wright (Eds.), Advances in consumer research Vol. 1. Urbana, Ill.: Association for Consumer Research, 1973. pp. 310-318.
- Wilkie, W. L. & Pessemier, E. A. Issues in marketing's use of multi-attribute attitude models. Journal of Marketing Research, 1973, 10, 428-441.
- Wilson, D. T., Mathews, H. L., & Monoky, J. F. Attitude as a predictor of behavior in a buyer-seller bargaining situation: An experimental approach. American Marketing Association Combined Conference Proceedings, Spring and Fall, 1972, Series No. 34. Chicago, Illinois: American Marketing Association, 1973, 390-395.





PRODUCT CLASS EXPERIENCE, DIMENSIONALITY AND RELIABILITY:  
THEIR RELATIONSHIP IN A NONMETRIC SCALING STUDY<sup>1</sup>

Adrian B. Ryans  
School of Business Administration  
The University of Western Ontario

Terry Deutscher  
College of Administrative Studies  
The Ohio State University

A method of estimating the number of dimensions a subject uses in making interbrand comparisons is presented. Using data from an experimental study on an inexpensive consumer durable, the relationship between an individual's product class experience and the number of dimensions used is examined. It is hypothesized that dimensionality and product experience are related in the following manner: a) individuals who received the product as a gift (i.e., they have usage experience but no shopping experience) use the fewest dimensions; b) individuals who have never owned the product (i.e., they have no usage experience and no shopping experience) use the most dimensions; and c) individuals who bought the product themselves (i.e., they have both shopping and usage experience) use an intermediate number of dimensions. It is further hypothesized that reliability in interbrand comparisons is inversely related to the complexity of the individual's perceptual structure.

### Introduction

Until the last two or three years, the task of determining the underlying dimensionality of a set of similarity data that had been submitted to a nonmetric multidimensional scaling program was a very subjective one. Recently a number of approaches to assist in this task (Wagenaar and Padmos, 1970; Isaac and Poor, 1974; Spence, 1972) have been proposed. A closely related but more important, question from the viewpoint of a marketing researcher using nonmetric multidimensional scaling is "What factors influence the number of dimensions a subject will use in making similarity judgments?" Previous researchers (Wilkie and Weinrich, 1972; Wilkie and Pessemier, 1973) have provided evidence of individual differences among consumers in the number of dimensions they use in making interbrand comparisons within a product class. One might further hypothesize that the people using the most dimensions ceteris paribus will be the least reliable in making their similarity judgments, because they have the most complicated frame of reference, and, consequently, the greatest chance of altering their basis of comparison from one point in time to another (Schroder, Driver, and Streufert, 1967). This also is an important factor that must be considered by the marketing researcher.

The objectives of the study reported here are to develop specific operational hypotheses on these questions and to test these hypotheses in the context of a particular nonmetric multidimensional scaling study undertaken by one of the authors. Following the analysis will be a discussion of the implication of the study for marketing professionals who are using nonmetric multidimensional scaling to aid in the understanding and prediction of consumer behavior.

## The Hypotheses

### Product Class Experience and Dimensionality

While it was expected that such a respondent characteristic as education would have some influence on the number of dimensions used in making brand comparisons, it was felt that the major influencing role would be played by the respondent's product class experience. It can be argued that a consumer's product class experience has two major components: first, the experience a consumer obtains by exposure to product advertising and brand displays as he goes through the search and purchase decision steps in the buying process (shopping experience), and second, the experience the owner of a brand acquires as he uses a brand in a product class (usage experience). But what effect will the amounts of these two components of product experience have on the dimensionality of a respondent's perceptual map? We would argue that for inexpensive consumer durable products at least, previous shopping experience tends to increase the number of dimensions used in making interbrand comparisons, while previous usage experience tends to reduce the number of dimensions. In fact, for this class of products, we will be willing to make a more specific set of hypotheses based on the two components of product class experience. These hypotheses will concern three sets of consumers--those who have never owned a brand in the product class, and two subsets of product class owners: those who bought the item themselves and those who received it as a gift.

Our hypotheses about the relative numbers of dimensions employed by each of these groups are based on the supposition that product usage enables a consumer to discriminate among a product class' different attributes according to their saliences. The net effect of this process is that, with increasing usage experience, a reduction occurs in the number of dimensions used in making interbrand comparisons. For example, a person who has never owned an electric blender might be able to find four or five attributes that differ from one brand to another when he first compares brands. However, after using a blender for a few months, he would realize that only one or two of these product class attributes really matter to him. For this reason, we will hypothesize that the people who use the most dimensions will be the ones who have never before owned an item in the product class.

Next, the mediating effect of shopping experience will be considered. The discussion here will focus on the two classes of owners of the product: those who bought the item themselves, and those who received it as a gift. The difference between the two groups lies in the fact that, although both classes have product usage experience, only the people who bought the item themselves have shopping experience. What difference will this experience make in the number of dimensions each person employs in making comparisons between apirs of brands?

Those who bought the item themselves have both types of product class experience, but they are generally not unbiased in the number of dimensions they use. That is, in shopping for and selecting a brand, they based their decisions on the different brands' relative standing on several dimensions--a number of dimensions that, according to our earlier supposition, is probably larger than the number that a large amount of product use experience would demonstrate to be important. Both the theory of cognitive dissonance (Festinger, 1957) and the theory of self-perception (Bem, 1970) would predict that such a respondent, having undergone this shopping experience, would subsequently use more dimensions in making interbrand comparisons than product use experience alone would suggest. In other words, the number of dimensions used

in subsequent decisions would be biased upward towards the number of dimensions the respondent used when making the purchase decision. On the other hand, the respondent who has received the product as a gift has not undergone shopping experience, and he will have no reason to subconsciously bias upward the number of dimensions used. Therefore, it seems reasonable to predict that the respondents who bought the item themselves will use more dimensions than the ones who obtained the item as a gift. To summarize, the results we would expect to obtain for all three classes of respondents are presented in Table 1.

In a number of consumer durable product classes, particularly inexpensive shopping goods, many of the features that differentiate the brands in the store turn out to have little practical value in actual usage of the product. It is this group of products for which the model developed here has the greatest applicability, and it would seem that the product class used in this study, electric blenders, clearly falls in this category. Thus we would hypothesize that the relationships represented in Table 1 would prevail.

#### Dimensionality and Reliability

It also seems reasonable to hypothesize that the more dimensions an individual uses in making similarity judgments between brands, the less reliable he will be in making those judgments. Used in this sense, reliability refers to how well the subject can replicate his first set of judgments at a second point in time. The hypothesis is based on the fact that, in general, the more dimensions an individual uses, the more likely it is that he will develop an integratively complex system for combining them (Schroder, Driver, and Streufert, 1967, p. 7). In this type of system, as opposed to an information processing system with a low integration index, an individual can use different methods for combining dimensions in making interbrand comparisons, rather than using one system for all situations. It seems plausible to expect, then, that the individuals using the least numbers of dimensions are the ones who provide the most reliable similarity judgments. Because these people have the simplest information processing systems, they are unlikely to alter their framework of reference in the course of providing the necessary paired similarity comparisons.

TABLE 1

Hypothesized Effect of Product Experience on the  
Dimensionality of Perceptual Judgments For  
Differing Amounts of Product Experience

Ever owned brand in the product class (usage experience)	Method of acquisition	Shopping experience	Number of dimensions used in making similarity judgments
No	--	No	Most
Yes	Bought by a member of the family	Yes	Intermediate
Yes	Gift	No	Fewest

### The Data

The data used to test the hypotheses developed above were gathered as part of a project to test a model of consumer choice behavior. The subjects in the experiment were a paid sample of 188 women recruited by a market research recruiting firm. Although the women were not selected randomly, they did represent a wide variety of socioeconomic backgrounds. The stimuli in the project were twelve brands of electric blenders, a broad cross-section of the blenders available in retail outlets on the San Francisco Peninsula in March, 1972. The blenders were displayed in the laboratory in a manner similar to the way they would be in the store, and the subjects had an opportunity to examine them thoroughly during the session. In the scenario which was used throughout the data gathering steps of each session, the subjects were told to imagine that they were buying a blender for their own use. If they already owned a blender they were told to imagine that it was broken and could not be economically repaired, and that they were therefore looking for a new one. To increase the realism of the situation, each subject was informed that several of the respondents would win one of the blenders and that the brand each of the winners would get depended on her choice behaviour in the study.

Among the data provided by 74 subjects in one experimental condition<sup>2</sup> were ratings of the overall similarity of all possible pairs of twelve blenders. The ratings were made on a nine-point rating scale anchored at "completely different" and "almost identical". In addition each subject also repeated the pairwise rating for four of the pairs of blenders (the same pairs were repeated in all questionnaires). The subjects were not informed that four pairs appeared twice in the list of seventy pairs of blenders, and none of them reported noticing the repetition. Seventy of the 74 subjects in this experimental condition provided sufficient data to be included in the analysis.

### The Methodology

In order to determine the number of dimensions each of the subjects was using in making her similarity judgments the individual similarity data matrices were first submitted to M-D-SCAL-V (Kruskal 1964a, 1964b), for scaling in five through one dimensions using Stress Formula One. The five stress values for each individual were then submitted to the M-SPACE program (Spence and Graef, 1973). M-SPACE is a program which determines the underlying dimensionality of an empirically obtained set of similarity data. The program is based on comparing the obtained set of stress values with those obtained in an extensive Monte Carlo simulation (Spence, 1970) in which the number of stimuli, the true underlying dimensionality, and the error level in the data varied.<sup>3</sup> An index of fit is computed by taking the root mean square deviation of the fitted from the obtained stress values. The dimensionality and the error level at which the fit index is minimized is then determined and this is considered to be the dimensionality and error level that best characterize the data.

Since the data for each individual consisted of sixty-six paired comparisons on a nine-point similarity scale, there were obviously many tied ratings. The results reported in the next section are those based on using the primary approach to handling ties in M-S-SCAL-V, since it appears to be more commonly used in marketing applications.<sup>4</sup>

Multiple regression was used to determine the relationship between the dimensionality obtained and such respondent characteristics as education and product class experience. Product class experience was represented by the square root of the number of times the respondent had used a blender in the

previous month, and dummy variables were used to represent non-owners (n=6), and owners who received it as a gift (n=34).<sup>5</sup>

A multiple regression analysis was also conducted to determine the relationship between dimensionality and reliability, while controlling on such variables as those representing product experience. The reliability with which a respondent was able to make the similarity judgments was determined by computing the Spearman rank correlation coefficient between the judgments for the four repeated pairs of blenders the first and second time they were made.

Ideally in both of the above sets of analyses we would like to have controlled on blender purchase and usage experience prior to obtaining the most recent blender. Unfortunately these data were not available. Since such experience would likely have occurred several years before this experiment, we do not think it would have a significant impact on the results reported here.

### The Results

The aggregated results of the analysis to determine the dimensionality each subject was using in making her similarity judgments are reported in Table 2.

TABLE 2

The Number of Subjects Using  
Different Dimensionalities In  
Making Their Similarity Judgments

Dimensionality	Number of subjects using the dimensionality
1	25
2	16
3	10
4	19
Total	70

The obtained dimensionality was then used as the dependent variable in the multiple regression analysis. Two models were used in the regression analysis, the first containing all the variables that it was hypothesized would influence the dependent variable, and the second containing only those that appeared to be most important. These results are reported in Table 3. The usage rate of the product class and the education of the respondent appeared to have no significant relationship with the number of dimensions that the respondent used, so these two variables were omitted in the second model. The product experience dummy variables both had the signs hypothesized, but only the variable representing owners who had received it as a gift was significant ( $p < .05$ , one-tailed test). That is, subjects who didn't own a blender tended to use about one-half

of a dimension more than subjects who owned blenders which they had bought themselves.<sup>6</sup> Subjects who owned a blender received as a gift used the fewest dimensions, approximately one-half of a dimension less than those in the other owner category.

The results of the regression analysis examining the factors influencing the reliability of respondent similarity judgments are presented in Table 4. The first model did not include dimensionality as an independent variable. The results of this regression indicated that gift owners were significantly more reliable than the other classes of respondents in making their similarity judgments. This is to be expected given the previous results which suggested they used fewer dimensions. The hypothesized intervening variable dimensionality was introduced on models (2) and (3). In (2), dimensionality has, as hypothesized, a negative coefficient ( $p < .01$ , one-tailed test), indicating that as the number of dimensions a subject uses rises, his reliability tends to decline. It is interesting to note from (2) that even after dimensionality is entered into the regression model, the gift owner dummy variable continues to have a significant positive coefficient. In (3), a simple regression of reliability on dimensionality, it is also clearly indicated that for each additional dimension a subject uses, his reliability tends to be .17 lower.

Before turning to the implications of this study, it is appropriate to comment briefly on the relatively low  $R^2$  values reported in Tables 3 and 4. As Morrison (1972) has pointed out, when one has a discrete dependent variable the  $R^2$  will be lowered somewhat, the amount depending on the true underlying distribution of the dependent variable and the number of discrete values the dependent variable is allowed to take. Therefore, the fact that dimensionality is a discrete variable with only four categories provides at least a partial explanation for the low  $R^2$  values obtained in Table 3.

The estimate of the reliability with which each subject is able to make similarity judgments is unlikely to be a very accurate one based as it is on only four repeated judgments. Ideally one would have liked to have a reliability measure based on a larger sample of repeated judgments. Thus there is likely to be a significant amount of error in the dependent variable in the regressions involving reliability. The effect of having this noise present in the observed reliability values is to reduce the observed  $R^2$ , and this is probably another factor contributing to the relatively low  $R^2$  values observed in Table 4.

#### The Implications

The results reported in the last section provide reasonable support for the hypotheses developed earlier. The effects of product class shopping and use experience on the number of dimensions that subjects used in comparing brands were in the hypothesized direction--blender owners who had received the brand as a gift tended to use about one less dimension than the respondents who did not own a blender. The analysis also suggested that the marginal impact of a subject using an additional dimension was a reduction of about .15 in her reliability (on a scale ranging from -1 to +1). Now that the hypotheses developed earlier do have some empirical support in one product class, we might tentatively suggest some practical implications for marketing managers and researchers using nonmetric multidimensional scaling.

Perhaps one of the most significant implications of the finding that product experience influences the number of dimensions used by a subject in making inter-brand comparisons is that product experience may provide a useful basis for

TABLE 3

Regression Analysis of Variables Influencing  
The Dimensionality Used

Model	Constant	Number of times blender used in previous month <sup>a</sup>	Education of respondent <sup>b</sup>	Product experience dummy variables		R <sup>2</sup>
				Does not own	Owner -gift	
(1)	2.10	-.02 (.23) <sup>c</sup>	.04 (.49)	.46 (.79)	-.51 (-1.67) <sup>d</sup>	.07
(2)	2.53			.47 (.87)	-.50 (-1.68) <sup>d</sup>	.07 <sup>e</sup>

<sup>a</sup>Square root transformation used.

<sup>b</sup>Years of schooling.

<sup>c</sup>Figures in parentheses refer to t values.

<sup>d</sup>p < .05 (one-tailed test).

<sup>e</sup>A test of the null hypothesis that R<sup>2</sup> = 0 can be rejected at the .10 level of significance (F statistic).

TABLE 4

Regression Analysis of Variables Influencing  
The Reliability of a Respondent's Judgments

Model	Constant	Product class experience dummy variables		Dimensionality	R <sup>2</sup>
		Does not own	Owner -gift		
(1)	.14	.07 (.27) <sup>a</sup>	.34 (2.48) <sup>b</sup>		.09 <sup>e</sup>
(2)	.51	.13 (.57)	.27 (1.99) <sup>c</sup>	-.15 (2.78) <sup>b</sup>	.18 <sup>d</sup>
(3)	.70			-.17 (-3.24) <sup>b</sup>	.13 <sup>d</sup>

<sup>a</sup>Figures in parentheses refer to t values.

<sup>b</sup>p < .01 (one-tailed test).

<sup>c</sup>p < .05 (one-tailed test).

<sup>d</sup>A test of the null hypothesis that R<sup>2</sup> = 0 can be rejected at the .01 level of significance.

<sup>e</sup>A test of the null hypothesis that R<sup>2</sup> = 0 can be rejected at the .05 level of significance.



aggregation. That is, if groups with varying amounts of product class experience do use different numbers of dimensions then it is appropriate to first split the sample on this variable before submitting the data to an aggregate non-metric multidimensional scaling analysis, or to a "points of view" analysis.<sup>7</sup> Segmenting the sample first on product experience has the practical advantage that it is a meaningful concept to marketing managers and it provides readily identifiable segments for developing specific communications plans. If the dimensions used by the various product experience segments are different or have markedly different saliences this may suggest that the evaluation criteria are quite different in the different segments. This again is useful information for managers charged with the task of developing an appropriate communications package. Furthermore, as a product class progresses through the product life cycle, the market changes from one where all respondents have little product class experience to a replacement market where practically all purchasers have at least some product class shopping or use experience. The company that fails to recognize this may be spending its communications budget very ineffectively.

It would seem that the results of this study tend to confirm what some marketing managers have long suspected--that different comparison mechanisms are likely to be employed by shoppers who differ in their amount of experience with a product class. First-time buyers of an inexpensive durable good are likely to be influenced by several attributes of the products. On the other hand, experienced users buying the item as a replacement, are more likely to have settled on a relatively few salient attributes that they desire. These findings have some important implications for a marketing manager's product design and promotional strategies.

The fact that the number of dimensions used by a respondent has a significant effect on the reliability of a subject's similarity judgments also has important implications for designers of studies using nonmetric multidimensional scaling. Clearly if this result is generalizable to other product classes it is an important factor that must be considered in determining the sample size for a study where aggregation of the data is to occur. For example, within a segment of respondents who share homogeneous perceptions of the product class, the level of error in the aggregate data (e.g., the means of each paired comparison judgment across the sample) is a function of the sample size and the reliability of the individual subjects. If the subjects in a segment are less reliable, more subjects must be aggregated to obtain a given level of precision in the aggregate estimate. In general, an improved knowledge of the factors influencing reliability should result in better sample designs by marketing researchers in nonmetric multidimensional scaling studies.

To this point we have said little about the generalizability of these results to other product classes. It is probably reasonable to hypothesize that similar relationships between product experience, dimensionality and reliability occur in similar electric appliance or inexpensive durable product classes. The relationship between dimensionality and reliability should be even more generalizable; it probably holds for most, if not all, marketing situations where nonmetric multidimensional scaling might be applied.

Finally we would suggest that in any marketing research study using non-metric multidimensional scaling it would be wise to conduct an analysis similar to the one described here either on the pretest data or on data from a small sample of subjects from the study. Such a study of product experience, dimensionality, and reliability, we contend would result in a better and more managerially useful analysis of the nonmetric multidimensional scaling data.

## FOOTNOTES

1. This research was supported in part by grants from the Marketing Science Institute and the Associates Research Fund, School of Business Administration, The University of Western Ontario.
2. This subgroup of respondents differed from the other experimental conditions by not being exposed to a new brand priced at one of three different levels. For a complete description of the data collection procedure, see Ryans (1973).
3. Similar simulation studies have been performed by others (Young, 1970; Wagenaar and Padmos, 1971; Sherman, 1973; Isaac and Poor, 1974).
4. In the primary approach no restrictions are placed on the fitted monotone regression values to a group of equal data values. In the secondary approach the fitted regression values are required to be equal when the original input data values are equal.
5. The square root transformation was used since it was felt that the marginal effect of each additional use would decline as the total usage increased.
6. This situation where subjects owned a blender and had bought it themselves is, of course, represented in the model by both dummy variables being zero.
7. A "points of view" analysis may be conducted by a Q-type principal component analysis of the type suggested by Tucker and Messick (1963) or by a clustering program such as Johnson's (1967) HICLUS program.
8. The results of this study also have practical implications for the attitude models outside the multidimensional scaling framework (for example, the brand preference models of Bass and Talarzyk (1972) and Bass, Pessemier, and Lehman (1973)), because they indicate that in making cognitive comparisons at least, shoppers with no previous product experience are likely to use more dimensions than previous owners. Measuring more dimensions than previous owners really use might lead to noise in the data.

## REFERENCES

- Bass, F. M., Pessemier, E. A., & Lehman, D. R. An experimental study of the relationships between attitudes, brand preference, and choice. Behavioral Science, 1973, 18.
- Bass, F. M., & Talarzyk, W. W. An attitude model for the study of brand preference. Journal of Marketing Research, 1972, 9, 93-96.
- Bem, D. Beliefs, attitudes and human affairs. Belmont, California: Wadsworth Publishing Company, 1970.
- Festinger, Leon. A theory of cognitive dissonance. Evanston, Illinois: Row, Peterson, 1957.
- Isaac, Paul D., & Poor, David. On the determination of appropriate dimensionality in data with error. Psychometrika, 1974, 39, 91-110.
- Johnson, S. C. Hierarchical clustering schemes. Psychometrika, 1967, 32, 241-254.

- Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a non-metric hypothesis. Psychometrika, 1964, 29, 1-27. (a)
- Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. Psychometrika, 1964, 29, 115-129. (b)
- Morrison, D. G. Upper bounds for correlations between binary outcomes and probabilistic predictions. Journal of the American Statistical Association, 1972, 67, 68-70.
- Ryans, A. B. Estimating consumer preferences for a new durable brand at a given price in an established product class--The development of a model and an experimental test. Unpublished doctoral dissertation, Stanford University, 1973.
- Schroder, Harold M., Driver, Michael J., and Streufert, Siegfried. Human information processing. New York: Holt, Rinehart, and Winston, 1967.
- Sherman, C. R. Nonmetric multidimensional scaling: A Monte Carlo study of the basic parameters. Psychometrika, 1972, 37, 323-355.
- Spence, Ian. Multidimensional scaling: An empirical and theoretical investigation. Unpublished doctoral dissertation, University of Toronto, 1970.
- Spence, Ian. An aid to the estimation of dimensionality in nonmetric multidimensional scaling. Unpublished manuscript, The University of Western Ontario, 1971.
- Spence, Ian, and Graef, Jed. The determination of the underlying dimensionality of an empirically obtained matrix of proximities. Multivariate Behavioral Research, 1974, 9, 331-341.
- Tucker, L. R., and Messick, S. An individual differences model for multidimensional scaling. Psychometrika, 1963, 28, 333-367.
- Wagenaar, W. A., and Padmos, P. Quantitative interpretation of stress in Kruskal's multidimensional scaling technique. British Journal of Mathematical and Statistical Psychology, 1971, 24, 101-110.
- Wilkie, William L., and Pessemier, Edgar A. Issues in marketing's use of multi-attribute attitude models. Journal of Marketing Research, 10, 428-441.
- Wilkie, William L., and Weinrich, Rolf P. Effects of the number and type of attributes included in an attitude model. In M. Venkatesan (Ed.), Proceedings of the 1972 Annual Conference of the Association of Consumer Research. Pp 325-340.
- Young, F. W. Nonmetric multidimensional scaling: Recovery of metric information. Psychometrika, 1970, 35, 455-473.

THE MODERATING EFFECT OF PRODUCT KNOWLEDGE  
ON MULTI-ATTRIBUTE ATTITUDE MODEL PREDICTIONS

Richard L. Oliver<sup>1</sup>  
University of Kentucky

The effect of product familiarity on instrumentality model predictions was assessed by adding "don't know" (DK) alternatives to the attribute instrumentality set of a new product attitude study. Subjects were classified into subsamples based on the number of DK responses and predictor-criterion correlations were then calculated for each subsample separately. Results obtained with two instrumentality models showed that the correlations were positively related to product knowledge although a reversal of this trend was observed at the high DK extreme. Further analysis indicated that number of DKs could be predicted from product familiarity variables and product specific confidence. The results imply that non-rational explanations for attitude formation may be required when the assumption of product familiarity is not tenable.

The predictive ability of various multiple attribute attitude models including those of Rosenberg (1956), Fishbein (1967), and others reported in Cohen, Fishbein, and Ahtola (1972) is fairly well established (Bass & Wilkie, 1973; Wilkie & Pessemier, 1973). In fact, the current status of these models is such that poor results can usually be traced to methodological problems (Fishbein, 1971). This has served to shift the academic interest in multi-attribute models from simple attempts to predict overall affect to discovering conceptual improvements that will increase the accuracy of the model's predictions (e.g., Bass & Wilkie, 1973; Craig & Ginter, 1974; Ahtola, 1973). One implication of this shift in emphasis is that poor results are interpreted as meaning that the researcher has mis-conceptualized his model or has committed a methodological error. The purpose of this paper is to demonstrate that low correlations between affect and summated instrumentality times value terms may result from factors other than concept or method.

The various multi-attribute attitude formulations used in consumer behavior are essentially rational models of consumer preference. Given a brand  $j$  and  $n$  salient attributes, a consumer is presumed to base his overall affect for that brand on the summated products of the degree to which brand  $j$  possesses attribute  $i$  (instrumentality) and his evaluation or rated importance of  $i$ . Thus the formulation is compensating in that a high or positive instrumentality times importance product for one attribute can "balance" a low or negative product for another attribute. An underlying assumption of rational choice is implicit in this formulation. A consumer who considers a brand to possess important attributes should not dislike

that brand nor should he prefer another brand that lacked or blocked these same attributes.

It appears critical, then, that the consumer be able to make instrumentality assessments. This presumes that he is aware or believes he is aware of the degree to which brands possess salient attributes. It follows that if he cannot respond to instrumentality questions, his brand preference will be based on factors other than the terms in the instrumentality times value formulation. Sheth (1973) has explicitly accounted for this possibility in a recent paper. Drawing on the work of Hull (1952) and Howard (1965), Sheth posited a habit or conditioning construct that would explain preference when instrumentality beliefs were not operative. Specifically, a consumer may be conditioned to respond to an object in terms of affect without making a rational assessment or mentally multiplying anything.

A second explanation for judgments of affect in the absence of underlying supporting beliefs may be found in the literature on personality response sets (Jackson & Messick, 1958; McGee, 1962). Jackson and Messick (1958) have distinguished between the content and personal style of an individual's response and argue that style will be pervasive when the subject cannot respond to the content of the question. Although a number of response styles have been identified including social desirability, acquiescence, and deviation, attempts to control for these influences will depend to some extent on knowing the subject's personality which, of course, is no mean task.

The potential impact of response style on attitude model predictions is not easily predicted and may have an effect opposite to that argued here *if* subjects respond to both affect and instrumentality on style. This will serve to artificially inflate the predictor-criterion correlations when no attitude structure may exist at all. However, when affect is based on style due to an inability to assess content and the subject is given an opportunity to indicate that he does not know the answers to instrumentality questions, low or zero correlations between affect and instrumentality predictions would be expected.

Unfortunately, information relating to the issues raised here cannot be obtained from previous studies of multi-attribute attitude predictions because researchers have generally neglected the possibility that subjects may not have been able to respond to instrumentality items.<sup>2</sup> This is not to say that researchers have avoided the issue; rather they have allowed the respondent to ambiguously interpret scale midpoints. To clarify this point, the following examples are presented to illustrate the more common operational measures of instrumentality.

*The semantic differential.* A number of studies (Sampson & Harris, 1970; Scott & Bennett, 1971; Churchill, 1972) have used the traditional 7-point bipolar adjective scale as shown below.

Bad taste : \_\_\_ : \_\_\_ : \_\_\_ : \_\_\_ : \_\_\_ : \_\_\_ : Good taste

Because no scale values are typically shown, the subject is permitted to interpret the mid-range category on an individual basis.

*The integer scale.* Other studies (Lehmann, 1971; Hughes & Guerrero, 1971; Sheth & Talarzyk, 1972) have used the following type of scale:

Low 1 2 3 4 5 6 7 High

Again, each respondent interprets the midpoint, category "4," on an individual basis. Some researchers (Sheth & Talarzyk, 1972) have used only six integers thereby forcing the subject to respond with a positive or negative evaluation.

In both types of scales above, no explicit provision is made for a "don't know" response. While some subjects may have interpreted the midpoint as such, it is this author's opinion that the "4" response would be used as an "average" or "neither" category. As failure to include a "don't know" alternative may result in an answer when none exists (Boyd & Westfall, 1972), it appears critical that the basis for these answers be known.

Failure to explicitly include a "don't know" alternative in the instrumentality set of multi-attribute attitude studies has a number of implications other than those previously discussed. First, subjects who do not have the information needed to respond to a particular instrumentality item may answer in a fashion consistent with their overall attitude or on the basis of the response pattern of earlier belief items. This phenomenon would be predicted by a number of cognitive consistency theories (McQuire, 1972) and will serve to artificially inflate the predictor-criterion relationship.

A second implication concerns the unknown consequences of comparing results across subjects who are basing their preferences on different numbers of different attributes. This issue reduces to one of combining subjects who determine their overall affect for an object on the basis of a large set of attributes with those who use one or two as though both groups of individuals were equivalent and homogeneous. With few exceptions (Churchill, 1972; Craig & Ginter, 1974), authors have assumed that all subjects use the same number of "salient" attributes.

The problem above is further compounded if instrumentality items are scored 1 to 7 rather than -3 to +3. If a respondent selects the scale midpoint to indicate "don't know," the latter scoring procedure will correctly preclude that attribute from contributing to the subject's aggregate instrumentality *times* value score. The former scoring procedure, however, implicitly includes that item as though it were scored "neither" or "average." There is more to this issue than is reported here; interested readers are referred to Schmidt (1973).

In an attempt to provide answers to some of the questions raised in the introduction, this study was conducted with two purposes in mind. First the effect of adding "don't know" alternatives to instrumentality items on predictions made by a multi-attribute attitude model was investigated. Second, an effort was made to show that the number of "don't know" responses could be predicted by two constructs, product familiarity and product specific confidence. It was hoped that this information would suggest a perspective for the role of product familiarity in models of consumer preference.

## Method

### Sample

Three groups of undergraduate students were asked to participate in a study of attitudes toward Ford's new Mustang II shortly after its introduction. The total sample was comprised of a "captive" classroom convenience sample (n=164), a "man-on-the-street" convenience sample (n=105), and a quota sample controlled for sex (n=142). Due to the subsample sizes needed in the analysis and because the distinctions between the three samples would not be expected to affect the results, the three groups were combined into a total sample of 405 usable replies. Sixty-three percent of the sample was male.

### Measures

In an attempt to reduce the effect of cognitive consistency, the variables were measured on a three-page questionnaire in the following manner. An index of overall affect, a self-rating of product specific confidence, and product familiarity questions were included on page 1. These measures were followed by attribute importance scales on page 2 and instrumentality items on page 3. It was hoped that positioning the importance scales between the affect and instrumentality measures would help to interfere with the subjects' recall processes and thus discourage them from responding to the belief measures on the basis of overall affect.

To obtain an evaluation of affect, subjects were asked to check how appealing the new Mustang II was to them on a seven point scale ranging from "Very unappealing" through "Neither" to "Very appealing." In a similar manner, product specific confidence was measured on a seven point scale ranging from "Very unconfident" to "Very confident." Finally, product familiarity was obtained from the following question:

Are you familiar with Ford's new Mustang II? Yes [ ] No [ ]

If yes, how do you know about it? (Check all that apply.)

- [ ] Have seen or heard ads for it.
- [ ] Have talked to someone that owns one.
- [ ] Have seen it in a dealership or on the street.
- [ ] Have ridden in or driven one.

Fifteen product attributes<sup>3</sup> were selected as important to students in car buying decisions from research summaries provided by Ford Motor Company and from an earlier study by Richmond, Kraft, and Hubbard (1973) on students' reactions to the Ford Pinto. Because all attributes were somewhat important (i.e., they would generally be rated as "good" on an evaluative dimension), attribute importance was measured on a relative basis (Scott & Bennett, 1971; Mazis & Klippel, 1973). That is, the subjects were asked to scale the importance of each attribute when compared to all others in an effort to discourage them from rating every attribute as important. Scale values ranged from "Not important" (0) through "Less important" (1) to "More important" (5).

Instrumentality items were constructed using a seven point bipolar adjective scale with the midpoint set apart from the two adjacent categories and explicitly labeled "Don't Know." Moreover, the scales were anchored at

the poles with good (+) and bad (-) adjectives so that no neutral term was used as a scale endpoint. Two typical instrumentality items are shown below:

	(-)	Don't Know	(+)	
Cramped	: ___ : ___ : ___ :	: ___ : ___ : ___ :	: ___ : ___ : ___ :	Roomy
Noisy	: ___ : ___ : ___ :	: ___ : ___ : ___ :	: ___ : ___ : ___ :	Quiet

Note that no "neither" category was included. If a student did not check "Don't Know," it was assumed that he could evaluate the Mustang II on that attribute as at least either slightly good or slightly bad.<sup>4</sup> The scale values were scored from -3 to +3. A "Don't Know" response was scored zero.

### Analysis

Subjects were placed in subsamples on the basis of the number of instrumentality "don't know" (DK) responses checked. Simple correlations were then calculated between the appeal scores and (a) summated instrumentality times importance terms and (b) summated instrumentality terms alone for each subsample of DK categories. These correlations were then treated as dependent variables (Jones, 1968) to determine if the subsamples could be drawn from a population with a common correlation.

To investigate the determinants of a "don't know" response set, number of DKs was used as a dependent variable and regressed on familiarity, confidence, and sex of the respondent. Because familiarity was measured dichotomously and by exposure category, two regressions were run. In the first, number of DKs was regressed on confidence, overall familiarity (yes = 1, no = 0) and sex (male = 1, female = 0). The second regression was similar to the first except that the single familiarity variable was replaced with the four exposure categories coded in dummy variable format.

### Results

The mean number of instrumentality DK responses over 15 attributes was 7.57. Thus, the average respondent was only knowledgeable in half of the attribute areas. This should immediately call to question the implicit assumption of product knowledge presumed in many consumer behavior studies. Of the 405 subjects, 51 (12.6%) checked "Don't Know" for every attribute. The remaining 354 subjects were distributed over the other DK categories with subsample *ns* ranging from 6 to 38.

To overcome the effect of sampling variations in the smaller DK categories, they were aggregated into groups of two. Respondents in the subsample who responded with a positive or negative evaluation for every instrumentality item (zero DKs) were considered unique and analyzed separately. The other seven subsamples included those with 1 and 2, 3 and 4, ..., and 13 and 14 DK responses. The group responding with 15 DKs would yield predictor-criterion correlations of zero and was not analyzed in this section of the study. The subsample *ns* are shown in Table 1 along with correlations between appeal and (a)  $\Sigma$  (instrumentality x importance) and (b)  $\Sigma$  instrumentality scores.



TABLE 1

Correlations Between Appeal and Two Instrumentality Models  
as a Function of the Number of Don't Know Responses

Don't Know Categories	n	$\Sigma$ (Instrumentality x Importance)	$\Sigma$ Instrumentality
0	20	.688	.686
1,2	36	.743	.749
3,4	56	.597	.578
5,6	68	.448	.519
7,8	60	.482	.435
9,10	63	.293	.349
11,12	36	.371	.300
13,14	15	.659	.684

It is apparent from the results that product knowledge clearly had a moderating effect on the results obtained with the two instrumentality models. When product knowledge was complete or almost complete (i.e., 0-4 DKs), the correlations obtained were on the order of .6 or higher. However as the number of DK responses increased beyond that point, a rather consistent decline in the magnitude of the correlations occurred for both models up to the last DK category where a reversal of this trend occurred. The high correlations obtained for the group of subjects who were knowledgeable in only one or two attribute areas may have been an artifact of the small subsample  $n$  or may indicate that, when only one or two pieces of information are known about an object, one's attitude is formed rather consistently with this limited information.

To determine if the correlations obtained could be interpreted as being from different populations (e.g., knowledgeable and unknowledgeable subjects), a  $\chi^2$  test of independence was calculated on the  $z$  transforms for both models (Jones, 1968; Snedecor & Cochran, 1967). With seven degrees of freedom, a  $\chi^2$  of 14.07 is needed for significance at the .05 level. Correlations for the instrumentality times importance and instrumentality models yielded  $\chi^2$  values of 13.50 and 13.21 respectively. While these values are significant at the .10 level, the reader may wish to draw his own conclusion as to the acceptance or rejection of a common population correlation.

In order to determine if the number of DK responses could be predicted, DK was used as a dependent variable and regressed on product specific confidence, product familiarity, and sex. The correlation matrix for this set of variables is shown in Table 2 where it can be seen that all independent measures were negatively correlated with the criterion ranging in magnitude from -.20 to -.35. Thus, the number of DK responses was greater for females, subjects who rated themselves as having little confidence in their ability to judge automobiles, and subjects who had had little or no exposure to the Mustang II.

TABLE 2  
Intercorrelations Between Number of Don't Know Responses  
and the Independent Variables (N=405)\*

	1	2	3	a	b	c	d
1. No. of DKs	-						
2. Confidence	-.22	-					
3. Familiarity	-.35	.27	-				
a. Media	-.20	.12	.64	-			
b. W-O-M	-.34	.10	.29	.09	-		
c. Observation	-.27	.23	.57	.29	.32	-	
d. Product use	-.32	.11	.20	.05	.67	.34	-
4. Sex (Male=1)	-.20	.39	.29	.22	.13	.25	.08

\*Due to the large sample size, correlations greater than .10 and .13 were significant at the .05 and .01 levels respectively.

Because the intercorrelations between independent variables were generally moderate, an attempt was made to improve the degree to which the number of DK responses could be predicted through multiple regression. To achieve this, the dependent variable was stepwise regressed on two models: (a) confidence, overall familiarity, and sex, and (b) confidence, the four familiarity categories, and sex. The results reported in Table 3 show the "best" regression equation using the "Maximum R<sup>2</sup> Improvement" technique (Barr & Goodnight, 1972).

TABLE 3  
Results Obtained when Number of Don't Know Responses  
Was Regressed on the Independent Variables: Two Models

Model	R	R <sup>2</sup>	Variables in Model	Standardized β-Weight	Significance of β-Weight
(a)	.363	.132	Familiarity	-.298	.01
			Confidence	-.144	.01
(b)	.435	.189	W-O-M	-.202	.01
			Confidence	-.167	.01
			Product Use	-.166	.01
			Media	-.151	.01

The results show that the two models explained 13% and 19% of the criterion variance respectively. Product familiarity was clearly the best overall predictor although confidence played an independent contributing role in both models. Sex did not contribute any variance beyond that already explained probably due to the fact that it was correlated with confidence. The standardized β-weights show that overall familiarity played a greater role in determining DK responses than did confidence although when

familiarity was viewed in terms of specific types of exposure, the  $\beta$ -weights for all significant predictors were similar.

Surprisingly three of the four familiarity variables made independent contributions to the second regression model. Word-of-mouth, product usage, and media exposure were negatively related to the number of DK responses although word-of-mouth appeared to have a slightly greater effect on the criterion. The remaining familiarity measure, product observation, was not included in the regression equation due to its high correlation with product usage. The fact that three of the four measures entered the model is significant in that it suggests that the influence of product information sources may be additive. It also implies that exposure to different kinds and types of information is necessary for full brand comprehension, a concept implicit in Howard and Sheth's (1969) taxonomy of significative, symbolic, and social inputs to the buyer decision process.

#### Discussion

The results of this study demonstrate that low predictor-criterion correlations obtained using multi-attribute instrumentality models may not result from poor conceptual work or methodological flaws. An alternative explanation, that of product or brand unfamiliarity, may be equally tenable. This effect has been concealed in previous studies because researchers have generally failed to include a "don't know" alternative in the instrumentality set.

Thus, indirect but preliminary support has been shown for an extended model of affect which incorporates response style tendencies or conditioning terms (Sheth, 1973). In the absence of complete or nearly complete product knowledge, consumers may respond to affect on the basis of an unspecified but habitually determined response set. While Sheth did not investigate the effect of conditioning in his study, future researchers are advised to more fully specify their models by including an operational measure of habit, particularly when the assumption of product familiarity across subjects is suspect. Howard (1965), for example, has drawn on the work of Hull (1952) and suggested that habit may be measured by the absolute number of reinforced purchases. Further elaboration on this point will ultimately lead to a discussion of stochastic learning models; interested readers are referred to Massy, Montgomery, and Morrisson (1970).

The attempt to predict product knowledge was somewhat encouraging. While the variance explained in the two regression models was not exceptionally high, the results clearly show that the number of DK responses to instrumentality items can be predicted from product familiarity and product specific confidence. Had better measures of the predictors been obtained, the results may have been more convincing. As the reader may recall, all familiarity measures were coded in dummy variable fashion. More precise measures of media recall (Lucas & Britt, 1963) or of the nature and content of word-of-mouth (Arndt, 1967), for example, may have improved upon the relationships obtained here. Similarly, a more construct-oriented approach to the measurement of product specific confidence such as that used in perceived risk investigations (Cox, 1967) may also have increased the significance of the findings.

Hopefully this preliminary study on the influence of product knowledge in attitude model predictions will encourage researchers to extend their research beyond the current vogue of rational cognitive multiplication models to include other factors explaining nonrational or habitual preference.

#### FOOTNOTES

1. Richard L. Oliver is Assistant Professor in the Department of Business Administration, University of Kentucky. He thanks Professor Jagdish N. Sheth of the University of Illinois for his helpful comments and the University of Kentucky Research Foundation for providing partial support for this research.
2. Mazis and Klippel (1973) provide a notable exception. They restricted their study to subjects who had used all brands in a product category. It will be suggested later in this paper, however, that usage alone provides only one dimension of product familiarity.
3. The attributes selected for the study were roominess, quietness, price, acceleration, safety, popularity, manufacturer's reputation, luxury, fuel economy, handling, comfort, styling, warranty, construction, and luggage space. Factors for which no information existed at the time of the study (e.g., maintenance cost, resale value) were purposely excluded.
4. To the extent that this was not true, the results may have been affected. Future researchers may wish to include both "don't know" and "neither" categories.

#### REFERENCES

- Ahtola, O. T. The vector model of preferences: An alternative to the Fishbein model. Faculty working paper, University of Florida, 1973.
- Arndt, J. Role of product-related conversations in the diffusion of a new product. Journal of Marketing Research, 1967, 4, 291-295.
- Barr, A. J. & Goodnight, J. H. Statistical analysis system. Raleigh, N.C.: North Carolina State University, 1972.
- Bass, F. M. & Wilkie, W. L. A comparative analysis of attitudinal predictions of brand preference. Journal of Marketing Research, 1973, 10, 262-269.
- Boyd, H. W., Jr. & Westfall, R. Marketing research. (3rd ed.) Homewood, Ill.: Richard D. Irwin, 1972.
- Churchill, G. A. Linear attitude models: A study of predictive ability. Journal of Marketing Research, 1972, 9, 423-426.
- Cohen, J. B., Fishbein, M. & Ahtola, O. T. The nature and uses of expectancy-value models in consumer attitude research. Journal of Marketing Research, 1972, 9, 456-460.

- Cox, D. F. Risk taking and information handling in consumer behavior. Cambridge, Mass.: Harvard University, 1967.
- Craig, C. S. & Ginter, J. L. An investigation of the effect of the dependent measure and attributes selected on the predictive power of the multi-attribute attitude model. Proceedings, Marketing Educators' Conference, American Marketing Association, 1974.
- Fishbein, M. A behavior theory approach to the relations between beliefs about an object and the attitude toward the object. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: Wiley, 1967. Pp. 389-399.
- Fishbein, M. Some comments on the use of "models" in advertising research. In Seminar on "Translating advanced advertising theories into research reality." Amsterdam: ESOMAR, 1971. Pp. 297-318.
- Howard, J. A. Learning and consumer behavior. In J. A. Howard (Ed.), Marketing theory. Boston, Mass.: Allyn and Bacon, 1965. Pp. 100-112.
- Howard, J. A. & Sheth, J. N. The theory of buyer behavior. New York: Wiley, 1969.
- Hughes, G. D. & Guerrero, J. L. Testing cognitive models through computer-controlled experiments. Journal of Marketing Research, 1971, 8, 291-297.
- Hull, C. L. A behavior system. New Haven, Conn.: Yale University Press, 1952.
- Jackson, D. N. & Messick, S. Content and style in personality assessment. Psychological Bulletin, 1958, 55, 243-252.
- Jones, M. B. Correlation as a dependent variable. Psychological Bulletin, 1970, 70, 69-72.
- Lehmann, D. R. Television show preference: Application of a choice model. Journal of Marketing Research, 1971, 8, 47-55.
- Lucas, D. B. & Britt, S. H. Measuring advertising effectiveness. New York: McGraw-Hill, 1963.
- Massy, W. F., Montgomery, D. B., & Morrison, D. G. Stochastic models of buying behavior. Cambridge, Mass.: M.I.T. Press, 1970.
- Mazis, M. B. & Klippel, R. E. Instrumentality theories and consumer attitudes: Comparing alternative models. Proceedings. Fourth Annual Conference, Association for Consumer Research, 1973.
- McGee, R. K. Response style as a personality variable: By what criterion? Psychological Bulletin, 1962, 59, 284-295.
- McQuire, W. J. The current status of cognitive consistency theories. In J. B. Cohen (Ed.), Behavioral science foundations of consumer behavior. New York: Free Press, 1972. Pp. 253-274.

- Richmond, D., Kraft, A., & Hubbard, C. Some attitudinal shifts of the automobile purchaser. Akron Business and Economic Review, 1973, 4 (2), 54-58.
- Rosenberg, M. J. Cognitive structure and attitudinal affect. Journal of Abnormal and Social Psychology, 1956, 53, 367-372.
- Sampson, P. & Harris, P. A user's guide to "Fishbein." Journal of the Market Research Society, 1970, 12, 145-166.
- Schmidt, F. L. Implications of a measurement problem for expectancy theory research. Organizational Behavior and Human Performance, 1973, 10, 243-251.
- Scott, J. E. & Bennett, P. D. Cognitive models of attitude structure: Value importance is important. In F. C. Allvine (Ed.), Relevance in marketing: Problems, research, action. Chicago: American Marketing Association, 1972. Pp. 346-350.
- Sheth, J. N. A field study of attitude structure and attitude-behavior relationship. Faculty working paper #116, University of Illinois, 1973.
- Sheth, J. N. & Talarzyk, W. W. Perceived instrumentality and value importance as determinants of attitudes. Journal of Marketing Research, 1972, 9, 6-9.
- Snedecor, G. W. & Cochran, W. G. Statistical methods (6th ed.). Ames, Iowa: Iowa State University Press, 1967.
- Wilkie, W. L. & Pessemier, E. A. Issues in marketing's use of multi-attribute attitude models. Journal of Marketing Research, 1973, 10, 428-441.



PERCEIVED INSTRUMENTALITY AND VALUE  
IMPORTANCE OF NEWSPAPER INFORMATION

Flemming Hansen  
Aalborg University,  
A.I.M. MARKEDSANALYSE a-s

In consumer behaviour much theory is formulated around value importance/perceived instrumentality variables. In the present paper such variables are studied in an untraditional context. In the paper changes in the functions performed by newspapers are analyzed. It is shown how the type of information carried, the role of entertainment and the ability to act as a social reference point for the reader has changed. Also it is discussed how the relative importance of these three major functions can be studied, and how the kind of audiences the papers attract depend upon how they mix these functions.

In addition to the "content factors" "treatment" factors are analyzed. The importance of understandability, readability, and optimal complexity is discussed, and it is shown how it varies depending upon the nature of the audience.

Based upon the above analysis, a model of the interactions between a publication and its audience is built. It describes how the audience's use of and satisfaction with the media varies with variations in "content" and "treatment". In support of the model it is shown how stable media "content" and "treatment" factors exist across different types of newspapers. Based upon factor analytical studies of consumers ratings of media a general factor structure is suggested.

Additionally a strong relationship is demonstrated between interest in different types of "content" and "treatment" factors and the papers which are being chosen, and it is discussed how different segments can be defined in terms of interest in and preferences for different combination of "content" and "treatment".

#### Instrumentality of Newspaper Information

The author has previously been concerned with the Rosenberg type of model in different areas of consumer behaviour.<sup>1</sup> Here this type of approach is applied in the area of information acquisition.

It has been proposed that newspapers serve three purposes for the reader. They may carry information which has:

1. Specific instrumentality, i.e. they carry information which the reader needs for specific purposes. This kind of information is attended to when specific information is needed. Examples of such information could be weather forecasts, stock market prices, television programmes, etc.



2. General instrumentality, i.e. information which in future situations may prove useful to the reader in more general ways. This kind of information is attended to, since the reader expects that it may be useful to him some time in the future. It can be information about political or other matters which may be of importance for the individual in future decision making or in future social relations, or it can be information about how important people look upon various issues. In the latter case the information is valuable, since the reader can use it as a reference point or as a guide line for his own opinions and behaviour.
  
3. Immediate instrumentality, i.e. information which is esteemed because of its immediate satisfaction of needs on part of the receiver. The reader turns to it "for the fun of it" without conscious or unconscious thought about its future relevance. Entertaining material belongs in this category.

It could be argued that any single item in a newspaper (article, story, etc.) performs all of the above three functions. However, as suggested by the examples, some items primarily serve one of the functions whereas other material primarily serves others.

Particularly the distinction is important between, on the one hand, the first two categories where the information has future relevance, and the last where the information is appreciated by the receiver because of its "entertaining effect".

Below these two functions will be discussed in some detail.

#### The Informational Function

Newspapers transmit information in society. Over the years the kind of information, which in the last two centuries has been carried by newspapers, has been transmitted in many different ways. The efficiency with which this information is distributed may be measured in different ways. One possibility is to look at the speed with which it is carried from its source of origin to those readers who demand it. Personal conversation, letter writing, speaking in front of large audiences, and the use of messengers, all are means of communication which were used in early days. With the invention of Gutenberg new ways of transmitting information came into use, and eventually in the 17th century the newspaper, as we know it, emerged.

Later improvements in communication techniques such as the telegraph, the use of steamboats and trains, etc. increased the speed with which information could be delivered from its source of origin to the newspaper and the speed with which the newspaper could be distributed from the place of printing to the receivers. Still, however, a century ago days could pass from the time of an event in, say, Paris, and the time at which information about the event was read in a Milano newspaper, which was distributed in, say, Verona.

The telephone and the telegraph improved the efficiency of this communication pattern, and so did changes in printing technique, but not until the emergence of radio and later television did anything exist which was a serious threat to the monopoly of the print media in the distribution of information.

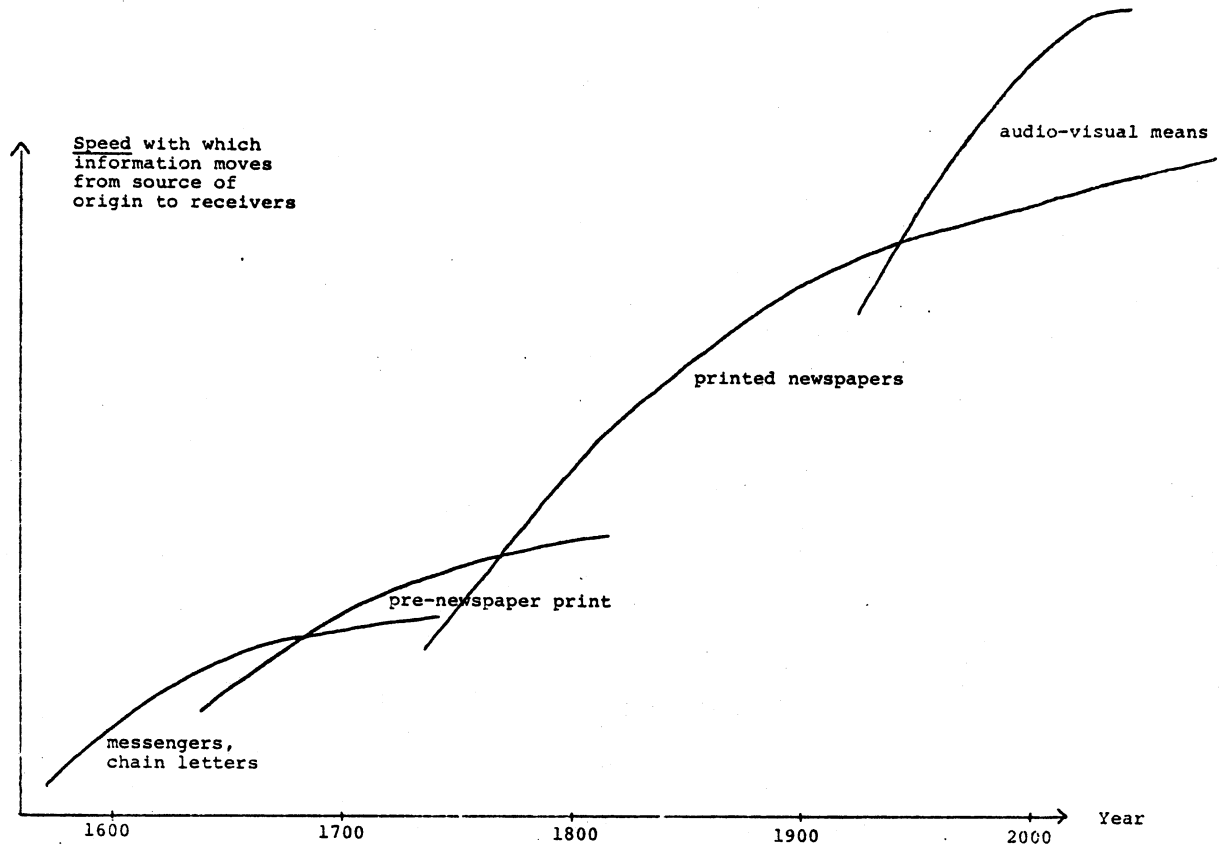


Figure 1. The development in the efficiency of mass communication.

In figure I. this development is illustrated. Different techniques have appeared at different times and for some period of time they have been the most efficient means of transmitting information. Then new techniques have come into use and changed the situation.

The overall picture illustrated in figure I. is typical for many development processes: An increasing growth rate explained by the emergence of new techniques and continuous improvements of these techniques. A similar pattern can be found in the development in the speed with which humans can travel, productivity of printing techniques, etc.

The fundamental suggestion made in figure I. is that in areas where speed of communication is vital, print media are loosing out to other media. Consequently print media must look for other needs which they can serve and they must look for types of information where factors other than absolute speed are important for the receivers. The first is done when the media more and more serve entertaining functions. This will be discussed subsequently. The latter is done when media either make the news themselves (as it is done by many popular dailies) or when they look for more specialized news which lends itself poorly for the large media: TV and radio.

Parallel with the development in the efficiency of communication, the amount of information to be transmitted has increased tremendously and many groups with different informational needs have emerged. Therefore, even though some media are superior with regard to general communication efficiency, it may still be possible to identify segments of readers with particular interests and to combine information which is relevant to such segments in such a way that newspapers or magazines doing this can be extremely efficient vehicles of communication for the receivers in question.

The extent to which particular media are able to identify homogeneous reader segments and to offer selections of information relevant to these segments is critical to the survival of these media. In the professional press and to some extent in the magazine press such a specialization has occurred. How such a segmentation strategy will have to look for dailies in the future is not obvious.

The dimensions along which this kind of segmentation should be carried out may be very different from those traditionally used in the study of newspaper readership. Rather than political views, geographical location, income, education, occupation, and similar variables, some yet unknown dimensions such as informational behaviour styles may have to be used. Below some of the research problems will be discussed which a newspaper encounters in the search for such variables.

#### The Entertainment Function

Realizing the growing competition when it comes to the transmission of instrumental information, magazines, and newspapers have become increasingly interested in entertainment and amusement material. Comics, popular stories, news with a heavy orientation towards entertainment, etc. have become more and more common. Also here, however, other means of communication have become serious competitors. The cinema, the radio, television, and in the future cable television and taped television will be important competitors.

Also here the most obvious opportunity for print media such as newspapers and magazines is to identify segments with such interests and background that they can be served more efficiently by the print media than by competing alternatives.

#### Form of Presentation

People vary in their ability to understand and absorb information. The extent to which the information is easily accepted depends upon factors such as its overall complexity. In turn this depends upon the use of pic-

tures, the readability of the copy, the sentence length, the type used, the nature of the material, etc. It is likely that any given reader<sup>2</sup> has an optimal level in any particular area. That is, within a given informational area an individual has an optimal level of complexity which he can accept. If the information becomes more complex he will tend to reject it and if it is less complex he also will refuse it.

Different such relationships exist for different individuals and even for the same individual different relationships exist in different context areas. In a particular area with which the individual is very familiar more complex material can be absorbed than in an area with which the individual is unfamiliar.

#### A Schematic Model

Let us summarize some of the points which so far have been made: Efficiency has to do with the speed with which relevant information is transmitted to motivated receivers in a form which is easily received (read, understood, and accepted). However, people vary widely in their informational needs and in their ability to receive information. One way of improving the efficiency with which a certain medium is communicating is to find out what is the ideal segment towards which the medium should be aimed and to identify what are the informational and other relevant characteristics of this segment.

How some of the factors which are likely to be important in a search for such segments interact is illustrated in figure 2. Here they are summarized in a model illustrating how they determine the efficiency of the communication.

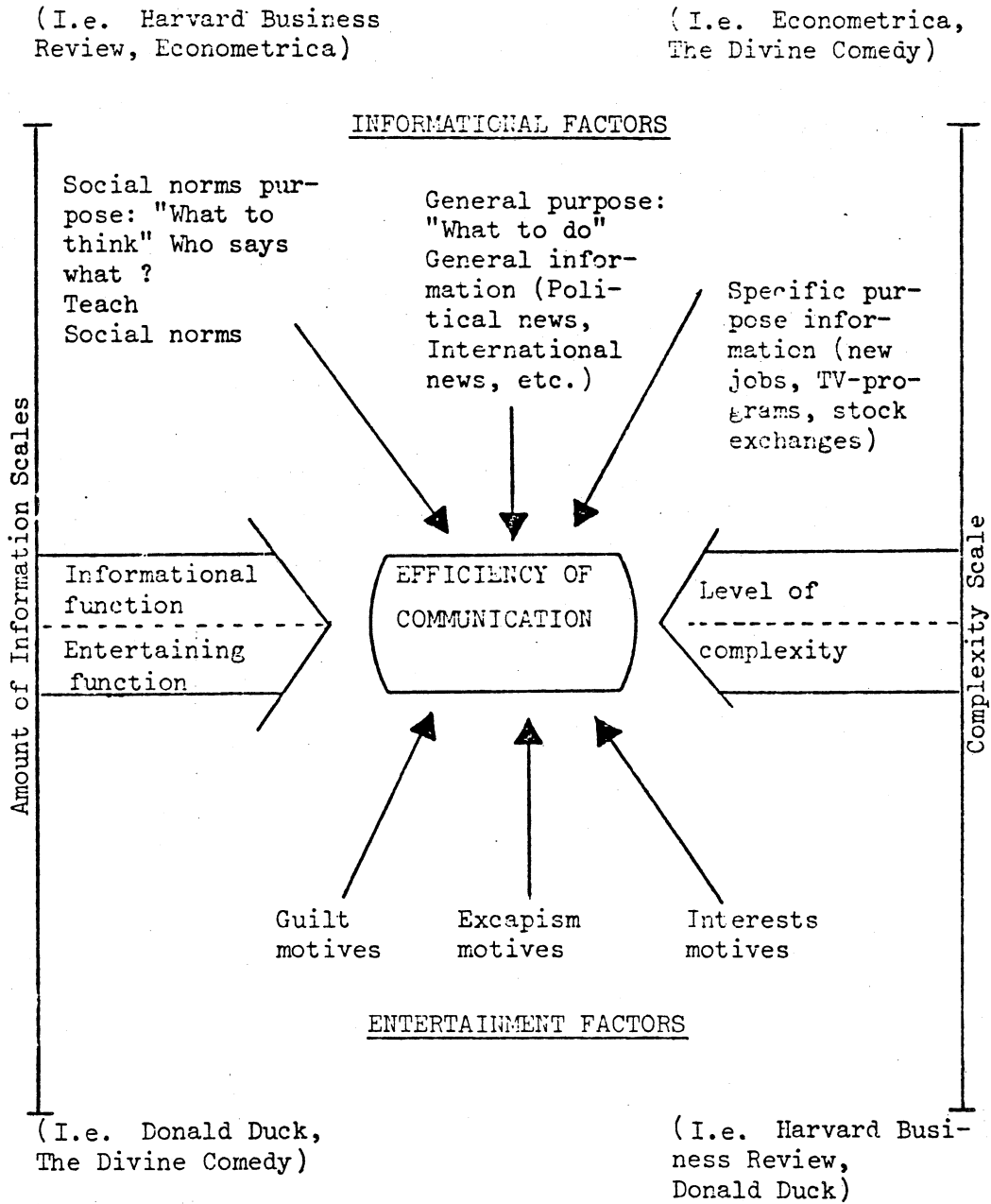


Figure 2. Content and treatment as determinants of the efficiency of communication.

In the upper part of the figure the informational factors are illustrated. In the lower part the amusement/entertainment factors are printed.

On the left hand side it is suggested how variations in amount of information may be seen as a summary measure of the content aspects.

On the right hand side a complexity scale is illustrated. This is meant to suggest how various presentation factors may be seen as different aspects of perceived complexity or rather as different aspects determining perceived complexity.

In the following pages, research will be discussed which has been aimed at identifying different informational ("content") and complexity or ("treatment") factors.

#### Informational (content) Factors

The research to be discussed has been carried out in Denmark and it deals with different Danish newspapers. For obvious reasons the identify of the papers being evaluated cannot be revealed. Also the factors and the averages being reported have been modified to make the data unidentifiable.

In search for the major informational factors newspaper readers see in daily newspapers, a list of informational categories was developed. In this way a total of 28 different news areas such as international news, theatre news, etc. were identified. For each category an "interest in the area" scale was formulated. This test battery was presented to approximately 400 newspaper readers who were asked to rate their daily newspaper. Based upon these ratings a factor analysis was carried out. An R-type analysis was carried out for each of four major newspapers. In these from 62 to 76 per cent. of the total variance has an explained. Since only little differences were found between the readers of the different papers a joint analysis was carried out. In this 68 per cent. of the total variance was accounted for.

Seven major content factors emerged from the 28 informational items which had been included in the study. These factors are illustrated in table I.

TABLE 1

## Seven Content Factors Derived from Factor Analysis

(Scales used: 1 = extremely interested,  
5 = not interested at all)

Number of respondents: 395.

	<u>Load</u>	<u>Average</u>
<u>FACTOR 1. "Intellectual interest factor"</u>		
Literary critics	0.83	3.519
Feature articles	0.75	2.924
Film and theatre reviews	0.73	3.223
The leader	0.67	2.771
<u>FACTOR 2. News Factor I. "Local or everyday factor"</u>		
Radio and TV-topics	-0.67	2.259
Good pictorial material	-0.63	2.341
News from Denmark	-0.55	1.799
Local news	-0.55	2.018
Material about prominent persons	-0.45	3.127
Foreign news	-0.45	2.771
<u>FACTOR 3. "Sport/hobby factor"</u>		
Motor matters	0.83	3.878
Sport	0.68	3.495
<u>FACTOR 4. "Consumption factor"</u>		
Fashion news	0.74	3.824
Classified advertisements	0.60	3.404
Textpage ads	0.61	3.670
Information about consumer goods	0.40	2.624
<u>FACTOR 5. News Factor II. "Social news Factor"</u>		
Information about social conditions	-0.75	2.003
Information about industrial relations	-0.63	2.805
Information about trade and industry	-0.49	3.086
Comments about politics	-0.47	2.574
<u>FACTOR 6. "Entertainment Factor"</u>		
Strip cartoons	0.79	3.213
Cartoons	-	3.101

	<u>Load</u>	<u>Average</u>
<u>FACTOR 7. "Reference group factor"</u>		
Letters to the editor	0.78	3.010
Personalia	.45	2.900
Material about prominent persons	.45	3.102

#### Form of Presentation Factors

Also for the form of presentation aspects a test battery was developed. With the use of this battery different newspapers were evaluated. A total of 84 items were included on the battery which was used to have 800 newspaper readers evaluate different newspapers. Subsequently these data were submitted to factor analysis. In this case, however, a quite large number of factors emerged all of which were not easily interpreted.<sup>3</sup> In table II, only 9 selected form of presentation factors are presented.

As suggested earlier these "form of presentation factors" are likely to be of a nature where "too much and too little" are bad things. For example, "to take a stand" may be positive to some readers but negative to other readers, and if it is so, it is extremely important to find out how the newspaper should look on this particular dimension.

TABLE II

#### Factor Analysis: Form of Presentation

	<u>Load</u>	<u>Average</u>	
		<u>Paper A.</u>	<u>Paper B.</u>
		N = 88	N = 84
<u>FACTOR I. COMPREHENDABILITY</u>			
Be easy to read	(.80)	1.62	1.26
Talk common people's language	(.70)	1.60	1.30
To be read without qualifications/ background	(.68)	1.88	1.51
Not use foreign words	(.51)	2.90	2.19
Have good lay-out	(.41)	1.34	1.26
<u>FACTOR II. READABILITY</u>			
Have short captions	(.63)	2.00	1.70
Be discreet	(.57)	3.01	2.92
Have short sentences	(.55)	2.23	2.04
Boil down the material	(.55)	3.20	3.00
<u>FACTOR III. ORIENTATION FACTOR</u>			
Easy to find what one is looking for	(.69)	1.32	1.23



		<u>Average</u>	
	<u>Load</u>	<u>Paper A.</u>	<u>Paper B.</u>
		N = 88	N = 84
<u>FACTOR IV. TO BE DETAILED</u>			
Be exhaustive/careful	(.69)	1.42	1.31
Be serious	(.66)	1.75	2.28
Follow-up matters to the end	(.53)	1.53	1.44
Be varied	(.48)	1.68	1.72
Treat subject detailed	(.39)	1.54	1.47
<u>FACTOR V. UNBIASED REPORTING</u>			
Make the readers able to understand what is going on in society	(.68)	1.32	1.20
Report statements as they were meant	(.63)	1.33	1.27
Has captions corresponding to the content	(.60)	1.21	1.24
Should rather be late with a piece of news than publish before being sure	(.47)	1.44	1.41
<u>FACTOR VI. SENSATIONAL</u>			
Be sensational	(.73)	3.65	2.67
Have sensational news	(.64)	2.98	2.05
Straight from the shoulder	(.53)	2.80	2.07
<u>FACTOR VII. TO TAKE A STAND</u>			
Make it abundantly clear what they mean	(.74)	1.85	1.24
Take own stand	(.65)	2.07	1.95
Have opinions and stick to them	(.44)	1.83	1.55
One should be sure of its stand	(.40)	1.99	1.77
<u>FACTOR VIII. LANGUAGE</u>			
Sober language	(.68)	1.62	2.00
Keep to the point	(.40)	1.77	1.97
Be neutral	( - )	2.51	2.10
<u>FACTOR IX. GIVE OPOSING VIEWS</u>			
Gives many different answers	(.72)	2.37	2.39
Express pros and cons	(.40)	1.46	1.54
Be all-round	(.35)	1.43	1.41

## Test of Factors

At this point in time it was felt that before segments were identified based upon the factors which were found, it was natural to test the extent to which the factors could explain differences in readership and in preferences for different papers. To do this several analysis were carried out. For example in table III. it is shown how people who primarily read national morning newspapers vary in interests from people who primarily reads national noon papers, who again vary from people who primarily read local newspapers. Seemingly people have different interests along the dimensions identified and the papers they choose vary with the combination of interests which they have.

TABLE III

Interest in Different Types of Content by Readers of Different Newspapers

	Read national morning papers N = 120	Read popular noon papers N = 180	Read local papers N = 300
1 = very most interested 5 = not at all interested			
<u>FACTOR 1. Intellectual interest</u> (Literary critics, feature articles, etc)	2.54	3.31	3.38
<u>FACTOR 2. Daily news</u> (Radio and TV-topics, news from Denmark, local news, etc.)	2.46	2.77	2.10
<u>FACTOR 3. Sport/hobby</u> (Motor matters, sport, etc.)	3.71	3.56	3.94
<u>FACTOR 4. Consumption</u> (Fashion news, classified advertisement columns, display ads, information about consumer goods)	3.47	3.48	3.27
<u>FACTOR 5. Social news and foreign news</u> (Information about social conditions, industrial relations, trade and industry, comments about politics, news from abroad)	2.47	2.72	2.49
<u>FACTOR 6. Humour</u> Strip cartoons	3.61	2.92	3.46

For the treatment factors a different kind of analysis was carried out. Here a number of questions from the original battery were selected in such a fashion that each of the major factors were represented. This was done by choosing a question from each factor with high loading on that factor. Subsequently these questions were used as independent variables in a stepwise regression analysis for each paper where a preference measure for the paper was used as the dependent variable. In this way the factors were identified which were of particular importance for the different papers. In general, high, multiple correlations were found. Also the factors explaining the variation in the dependent variable were different from newspaper to newspaper. For example for one of the newspapers the following regression equation was established.

TABLE IV

Relationships Between Preferences for Paper "I" and Various Treatment Factors

$$X_i = 2.95 + 0.27 (\text{sensational}) - 0.27 (\text{to use a clean language}) \\ + 0.13 (\text{to entertain}) - 0.11 (\text{to be neutral})$$

(X = preference for paper "I")

It will be seen that this particular paper gets readers who value a sensational approach, who do not care much about the language used, who want an amusing newspaper and who want a paper which takes a stand. That the paper is a popular noon paper can be read directly from the table.

#### Segmentation

At this point it was felt that meaningful content and treatment variables were identified. The next problem was to identify homogeneous segments which were useful target groups for the papers for which the research was carried out. Here two approaches have been discussed. In the first, departure was taken in the existing market situation. That is, it was analyzed how people were changing between the major papers involved in the study.

Subsequently it was analyzed to which extent the treatment factors and the content factors which were identified could be used to explain these changes. For each of the major flows of readers it was analyzed how people moving from one to another newspaper differed in evaluation of treatment factors. For example for readers switching between two papers the results shown in table V. emerged.

TABLE V

Differences in Evaluation of Paper A. Between People Moving  
From Paper A. to B. and People Moving From B. to A.  
(the lower value the more paper A. is seem to have the charac-  
teristics in question).

	<u>From A. to B.</u>	<u>From B. to A.</u>
	N = 36	N = 42
<u>FACTOR I. Comprehendability</u>		
Intelligible language	1.41	1.46
Too many foreign words	5.44	5.92
Good lay-out	2.00	2.23
<u>FACTOR II. Social Criticism</u>		
An aggressive newspaper	4.00	3.69
<u>FACTOR III. Seriousness</u>		
A profound paper	3.44	3.92
<u>FACTOR IV. Firm Attitude</u>		
Firm attitude in politics	3.67	4.31
Firm opinions about matters it dealt with	2.94	3.92
Too vague attitudes	3.50	2.08
Change attitude too often	4.11	2.77
<u>FACTOR V. Sensation</u>		
Is sensational	3.28	2.23
<u>FACTOR VI. From all Angels</u>		
Elucidate the material from many points of view and state for and against	2.29	4.00

Here it can be seen how those readers who move from paper A. to the competing paper differ from those who switch from paper B. back to A. They do so in the extent to which they want a newspaper which takes a clear position on the issues it is covered, in the extent they want the paper to be sensational, and in the extent to which they want their paper to be presenting pro et cons. If A. wants to gain readers from B. these are the dimensions on which the approach should build. Moreover, by carrying out analysis of this nature for the different competitive relationships it was possible to estimate what would be the effect of changing the existing newspaper in one direction or another.

The second segmentation approach would depart far away from the existing newspapers. Rather the question would be asked: "To what an extent is it possible to identify groups of readers with homogeneous interests in content and treatment"?

The technique to be used would be backwards segmentation departing in evaluation of interest in different content and treatment aspects. The results of this part of the project are not available, but it is hoped that it would identify groups of newspaper readers with a homogeneous set of interests which better could be served by a modified or a new product than whatever papers they are presently reading.

The extent to which this approach would provide successful results will depend upon the measurement devices used and the extent to which the available papers have been successful in adopting ideally to the existing market. Whether the latter is the case or not it would be possible to see from the study since it will be possible to analyze for each of the segments whether the readers in that segment tend to read only one or two newspapers or whether they have divided their newspaper reading among a variety of different papers. If the latter is the case a product adopted to the needs and interests of the people in that particular segment may have a chance.

#### FOOTNOTES

1. Flemming Hansen, "Consumer Choice Behavior - A Cognitive Theory", The Free Press, New York, 1972.
2. D.E. Berlyne, "Conflict, Arousal and Curiosity", New York, 1960.
3. With a total of 72 per cent. of the variance explained, 19 factors emerged. The unclear nature of some of these may possibly be ascribable to the test battery including several other than form of presentation statements. Here only a selection of the factors are discussed.

IDEAL POINT VERSUS ATTRIBUTE MODELS OF BRAND PREFERENCE :  
A COMPARISON OF PREDICTIVE VALIDITY<sup>1</sup>

Bernard Dubois<sup>2</sup>

Centre d'Enseignement Supérieur des Affaires  
Jouy en Josas, France

This paper reports and discusses the results of a comparative analysis of the predictive power of three classes of brand preference models (ideal point models, attribute models, and integrative models) in relation to an educational service. Results obtained at the group and individual levels demonstrate the superior performance of ideal point models, followed by integrative models. Also, the predictive value of structurally simple models is found to be comparable to that of more complex formulations. Finally, the increase in performance gained by giving explicit consideration to attribute weights is marginal. The paper discusses the conceptual, methodological, and managerial implications attached to these findings.

#### Introduction

The ultimate purpose of marketing activity is to influence buying behavior. To induce such behavior, marketers must understand how consumers generate their decisions and the role played by marketing stimuli in the decision-making process. Buying behavior, however, is so complex that the relationships existing between marketing stimuli and purchase activity are unlikely to be simple and direct ; consumers are under the influence of a host of mediating factors or situational variables such as social norms, word-of-mouth, financial status and the like.

Confronted with this situation, marketers have often found it useful to turn their attention to various pre-purchase stages, hoping that for these stages it would be somewhat easier to assess and predict the performance of particular marketing actions. In deciding which stage would be most appropriate, they have generally considered two specific but somewhat contradictory criteria : the selected stage should not be too remote from the purchase act so as to preserve the connection with sales, and yet relatively "pure" in order to facilitate the identification of marketing stimulus influence.

Given these requirements, the stage of brand preference, defined as:  
An affective stage of the individual which reflects his relative valuation of the alternative brands he considers when contemplating the purchase of a product of service

has often been considered an appropriate compromise. Sufficiently close to the purchase act to take into account the competitive nature of the buying situation, it seems remote enough to be largely unaffected by situational factors which inevitably blur the impact of specific marketing stimuli.

As a result, marketers have expressed a special interest in models that would explain and predict consumers' brand preferences. Among the various perspectives which have been adopted in the analysis of brand preferences, one of the most significant certainly is the information processing approach. More comprehensive than either subject or object-centered theories (that is, theories which exclusively focus on either consumer or product characteristics as explanatory variables), the information processing approach investigates the interaction existing between the subject and the object in terms of the mechanism whereby consumers evaluate the various alternatives available to them on the basis of a set of criteria reflecting their value orientations.

Although developed in so varied disciplines as attitude research,<sup>3</sup> decision theory,<sup>4</sup> micro-economics,<sup>5</sup> and psychometrics,<sup>6</sup> information processing models of brand preference may be classified into two major categories depending upon their specific postulates regarding the nature of the evaluative process: (1) Ideal point models, which rest on the assumption that individuals develop an ideal representation of the product they consider for purchase and then evaluate existing alternatives on the basis of their overall proximity to the ideal brand; and (2) attribute models which explicitly consider the desirability attached to each relevant attribute and explain preferences on the basis of the relative standing of alternatives along these attributes. More recently, a third class of models, integrative models, has been proposed as a compromise between the previous two categories. In these models, the ideal brand is decomposed into a set of ideal positions on each of the relevant attributes. Preferences are then understood in terms of attribute-specific deviations from the ideal.

While previous research has investigated the predictive power of each of these models considered separately, and while comparative research has been mostly limited to intra-rather than inter-class models, no researcher has yet tested on a single data base the relative performances of these three classes of models. Yet such a test is needed if one wants to discover which type of model is for a given type of product the best descriptor of consumers' brand preference structure. The objective of this research is to conduct and interpret such a comparative analysis.

#### Method

The relative performances of alternative models can only be assessed in reference to one or several specific criteria. In the present study, brand preference models are compared on the basis of their predictive ability. Although by no means the only criterion (diagnostic power may be important too), it is undoubtedly the most important since few researchers would pay attention to a model with less than acceptable predictive power.

In order to test such power, a control set of preferences was

obtained by direct measurement and the degree of convergence existing between these preferences and those predicted by each model was investigated and used as a measure of performance. Comparison was made both at the ordinal and interval levels so as to discriminate between various finer versions of the selected models.

The product class under investigation in this research is the American Master's degree program in the field of business administration. Accordingly, the "consumers" of the product are graduate business students and the various "brands" correspond to the various American graduate schools of business. The choice of this specific environment was primarily guided by methodological considerations. It was hypothesized that the decision to enroll in a graduate business school was of sufficient perceived importance to facilitate the emergence and therefore measurability of significant brand preferences. A sample of one hundred students, all enrolled in the M.B.A program at Northwestern University, were asked to describe their preferences for seven graduate business schools : Harvard, Wharton, Chicago, Northwestern, Stanford, Michigan, and Indiana. These institutions were found to be schools most frequently mentioned in a series of unaided recall tests conducted with the same population. The fact that all respondents belonged to the same institution was not a problem since wide variation in preferences was observed.

#### Control Preferences

Rank-ordered preferences were obtained from respondents simply by asking them to indicate which school they would have selected first, second, ... if admitted to all schools mentioned above. While indirect in nature, this approach to preference measurement is nevertheless adequate given the purposes of the present research. It is certainly true the fact of being enrolled in a specific program influences school preferences ; however, it equally affects school perceptions and images, that is, model input data, and therefore does not "contaminate" the measure of convergence provided enough variance in preferences is available, which was the case in this study. Interval measures of preference were obtained by means of the constant sum method (Meftessel, 1947 ; Comrey, 1950) and processed according to the Torgerson algorithm (1958) to generate an interval scale on which preference scores for each stimulus could be located.

#### Ideal point models

School preference predictions based on the ideal point model were obtained by means of the nonmetric multidimensional scaling procedure ; respondents were asked to compare similarity triads involving the seven schools listed above plus an eighth hypothetical Ideal. From these data, multidimensional maps were obtained in which the existing schools and the ideal were conjointly represented and in which distances between the ideal and each existing school could be interpreted as preferences.

Three specific types of distances were tested corresponding to different assumptions regarding the shape of the utility curve implicitly used in the derivation of preferences (at the interval level) : simple euclidean distances, which correspond to a linear decrease in utility (Figure 1) ; exponential distances, which correspond to a sharp decrease in utility as soon as one moves away from the ideal (Figure 2) ;



and parabolic distances, which correspond to an inverted U-shaped utility curve (Figure 3).

#### Ideal point models

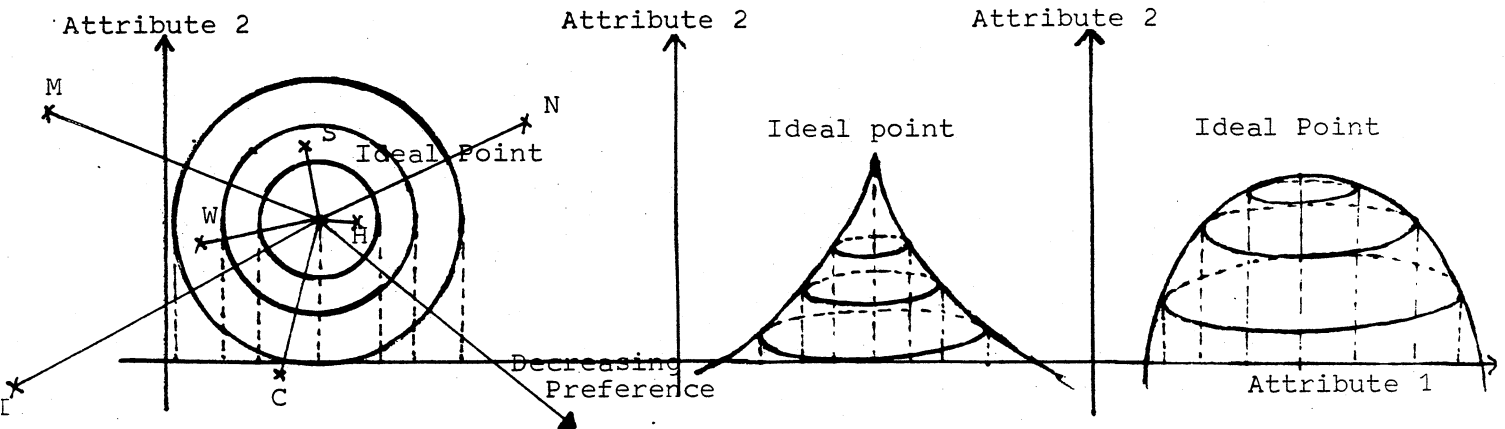


Figure 1:  
euclidean function.

Figure 2:  
exponential function.

Figure 3:  
parabolic function.

Although past research has mostly relied upon the euclidean model, several authors (Einhorn and Gonedes, 1971) have recently suggested that other forms of utility functions may lead to better descriptions of preferential judgments.

#### Attribute models

Three major types of attribute models were considered in this research : the single-attribute model, the lexicographic model, and the expectancy-value model. The raw data used to test these models consisted of a set of semantic differential scales developed from Osgood's lists (Osgood, Suci, and Tannenbaum, 1957), an analysis of schools' catalogs and advertising literature, and a series of in-depth interviews conducted with Northwestern M. B. A. students. The original list of scales was then factor-analyzed and reduced to twenty attributes which represented the most important and reasonably independent dimensions. The final set is reproduced in Table 1. Accompanying instructions were borrowed from Osgood et al. (1957, 81). The only modification was that, in addition to completing scales for each school including the ideal, each respondent was asked to indicate the relative importance of each semantic attribute (on a four-point scale).

The single-attribute model. Predictions based upon the single attribute model according to which all brands are perceived and evaluated on the basis of a unique dimension, were obtained as follows : for each subject (or the average subject in case of group analysis), a search procedure was initiated to discover the attribute considered the most important (on the basis of weights provided by respondents). The position of the ideal school on this attribute determined which side was the positive one, and the location of the existing schools resulted in the preference structure.

TABLE 1

Final Set of Semantic Scales

Quantitative vs Non-quantitative  
 Expensive vs Inexpensive  
 Unknown vs Well known  
 Theoretical vs Practical  
 Hard vs Easy  
 Teaching oriented vs Research oriented  
 Urban vs Non-urban  
 Passive vs Active  
 General vs Specialized  
 Conservative vs Liberal  
 Old vs Modern  
 Innovative vs Non-innovative  
 Small vs Large  
 High prestige vs Low prestige  
 Formal vs Informal  
 Involved vs Indifferent  
 Domestic vs International  
 Influential vs Non-influential  
 Valuable vs Ordinary  
 Close by vs Far away

The lexicographic model - In its purest form, this model is a straightforward extension of the single-attribute model. According to the regular lexicographic model, the individual rank orders brands on the basis of their values on the attribute most important to him and, if two or more brands are tied on that dimension, breaks the tie by considering the second most important attribute and additional ones if necessary until a complete set of preferences is generated. Despite its simplicity, the regular lexicographic model has received some support from empirical studies of the consumer decision-making mechanism (Alexis, Haines, and Simon, 1968 ; Clarkson, 1963 ; Bettman, 1969 ; and Russ, 1970). Recently, a more sophisticated version of this model, the semi-order lexicographic structure, has been proposed in the literature and seems to perform better than the regular version (Russ, 1971). In this model, the second most important attribute is considered not only if the values obtained for two or more stimuli on the most important attribute are exactly identical but for all cases in which the differences between these values are considered non-significant or non-noticeable. In the present research, preference predictions were generated from both the regular and semi-order versions of the lexicographic model.

The expectancy-value model. As opposed to the lexicographic model, which is non-compensatory in nature, the expectancy-value model, which seems to enjoy some popularity among marketing researchers, allows low scores on certain attributes to be compensated by high scores on others. This is made possible by its algebraic structure which is of the form :

$$Y_j = \sum_{i=1}^n w_i B_{j,i}$$

where :

$Y_j$  = index of attractiveness attached to brand j  
 $w_i$  = weight attached to attribute i  
 $B_{j,i}$  = value of brand j on attribute i.

As applied to the analysis of brand preferences, this model, which finds its roots in the work of attitude researchers such as Rosenberg (1956) and Fishbein (1967), has been found of good predictive value (Bass and Talarzyk, 1969 ; Hansen, 1969). Some dispute (Sheth and Talarzyk, 1972) exists, however, as to whether the two components of the model are equally important or in fact are both needed. In the present study, both unweighted and weighted formulations of the expectancy-value model were tested and compared on the basis of their predictive power. The full set of twenty scales was used in the derivation of preference indexes.

### Integrative models

Preference predictions obtained from integrative models were generated as follows : Instead of simply using the ideal school's position on attributes as an indicator of which side was favorable, the specific location of the ideal school was explicitly taken into consideration. The main difference between predictions based on expectancy-value and integrative models was that in the latter deviations from the ideal instead of raw scores were used.

The most general functional form of the integrative model can be represented as follows :

$$Y_j = \left[ \sum_{i=1}^n w_i (B_{j,i} - B_i^*)^r \right]^{1/r}$$

where :

- $Y_j$  = distance from existing brand  $j$  to the ideal brand (measure of preference)
- $w_i$  = weight attached to attribute  $i$
- $B_{j,i}$  = value of brand  $j$  on attribute  $i$
- $B_i^*$  = value of the ideal brand on attribute  $i$
- $r$  = Minkowski order ( $r > 1$ )
- $n$  = number of attributes

Thus far, this model has received limited testing, and the only two studies which have analyzed it (Lehmann, 1971 ; Bass, Pessemier, and Lehmann, 1972) have generated disappointing results. Yet, the model has high intuitive appeal. In this study, four types of predictions were obtained from the model depending upon whether weights were or not considered and whether euclidean ( $r=2$ ) or city-block ( $r=1$ ) distances were used.

### Empirical Findings

Data will be analyzed both at the group and individual levels. The essential reason for presenting individual results is theoretical : All brand preference models considered in the present study have been proposed to explain individual rather than group choice behavior. Yet, managers of educational programs may express limited interest in data displayed at the level of each student, given that they are not generally in a position to tailor their offering to individual needs. For this reason, results will be presented first at the group level. Of course, the assumption of homogeneity of preference will be made in order to justify

the aggregation process. For example, in testing ideal point models, only group ideal points, as derived from aggregated data, will be considered. This assumption will be relaxed when individual results are shown.

### Group level results

Tables 2 and 3 exhibit the overall brand preference structure revealed at the group level. It is against the values presented in these tables that preference model predictions were tested.

TABLE 2  
Overall Brand Preference Structure :  
Ordinal Level

	ranked 1st	ranked 2nd	ranked 3rd	ranked 4th	ranked 5th	ranked 6th	ranked 7th	final rank
Harvard	41	27	12	8	6	1	5	1
Stanford	31	26	17	15	5	3	3	2
Northwestern	23	17	36	14	8	0	2	3
Wharton	3	16	17	23	16	16	9	4
Chicago	1	8	11	16	35	20	9	5
Michigan	-	5	7	17	21	36	14	6
Indiana	1	1	-	7	9	24	58	7

Note : Circled values denote modal ranks. The data presented in this table tend to support the hypothesis of intersubject homogeneity of preference

TABLE 3  
Overall Brand Preference Structure  
Interval Level

Harvard	Stanford	Northwestern	Wharton	Chicago	Michigan	Indiana
2.338	1.898	1.608	0.913	0.721	0.608	0.351

It should be noted that the results presented in Tables 2 and 3 are perfectly consistent thereby enhancing the validity of the constant-sum method as a preference measurement instrument.

Table 4 exhibits, in summary form, the comparative performances of all structural models considered in the present research. Four main conclusions may be derived from an analysis of this table :

1. The ideal point model performs consistently better than all other formulations. This is especially true for the exponential version of the model which exhibits an impressive predictive power (almost 90 % of variance explained).

2. The supremacy of ideal point formulations is further confirmed by the fact that the city-block version of the integrative model is found to be superior in predictive ability to the expectancy-value model, while the only difference between the two structures is that explicit consideration is given to ideal positions on semantic attributes in the former model.

TABLE 4

Comparative analysis of predictive power of  
selected brand preference models :  
Aggregate level.

Brand preference models	Rank-order Correlation with ordinal overall preferences	Product-moment correlation with interval overall preferences
1. Ideal point models		
1.a. Euclidean formulation	1.00**	.9459**
1.b. Exponential formulation		.9721**
1.c. Parabolic formulation		.8317**
2. Attribute models		.767**
2.a. Single attribute model	.893**	
2.b. Lexicographic model		
2 b i Regular version	.893**	
2 b ii Semi-order version		
.05 threshold	.857**	
.10 threshold	.893**	
.25 threshold	.893**	
.50 threshold	.810*	
2.c. Expectancy-value model		
2 ci. unweighted formulation	.6429	.6436
2 cii weighted formulation	.6429	.6796*
3. Integrative models		
3 a. Unweighted city-block for- mulation	.7500*	.6893*
3.b. Unweighted euclidean formu- lation	.3214	.5463
3.c. Weighted city-block formulation	.7500*	.7551
3.d. Weighted euclidean formulation	.8571**	.6096

\*\* P < .01

\* P < .05

3. Students seem to need only a few dimensions to generate their complete set of preferences. This statement is suggested by the good performance of structurally simple models such as the single-attribute and lexicographic models. Conversely, when many attributes are considered, as in the expectancy-value and integrative models, it seems that a considerable amount of "noise" is added, either due to redundant or irrelevant information (or both). As a result, the predictive power of these models decreases.

4. Finally, weights seem to add little, if anything, to the predictive power of unweighted models.

#### Individual level results

Table 5 presents, in summary form, the respective performances of the various models tested in this research. Only ordinal correlations are displayed in this table given that, at the group level, ordinal and

interval coefficients were found to be highly congruent. Also, only one version of the lexicographic model is considered since little difference was observed between the various versions tested at the group level.

On the basis of the data exhibited in this table, it can be easily seen that the conclusions developed at the aggregate level also hold for individuals. The ideal point model is still found to be the most powerful model, at least from a predictive standpoint, its superiority being even more clearly established than previously. Then, adopting the number of correlation coefficients greater than .90 as a measure of predictive power, one successively finds : the single-attribute and lexicographic models which exhibit fairly similar performances ; the city-block version of the integrative model ; and, finally, the two versions of the expectancy-value model and the remaining formulations of the integrative model. As noted at the group level, it seems that weights do not generate additional predictive power and that a large number of attributes is somewhat detrimental to overall performance.

TABLE 5  
Comparative Analysis of Predictive power of  
Selected Brand Preference Models :  
Individual Level

	Greater than .9 but less than :	Rank-order correlation coefficients								
		.8 .9	.7 .8	.6 .7	.5 .6	.4 .5	.3 .4	.2 .3	.1 .2	.0 .1
1. Ideal point model	46	20	10	4	4	4	4	1	4	1
2. Attribute Evaluation Models										
2.a. Single attribute model	12	18	17	8	8	5	6	6	4	4
2.b. Lexicographic model (regular version)	13	17	15	11	9	4	5	7	6	3
2.c. Expectancy-value models										
2.c.i Unweighted formulation	7	23	9	17	14	6	9	3	2	2
2.c.ii. Weighted formulation	7	19	22	16	11	8	5	1	2	1
3. Integrative model										
3.a. Unweighted city-block formu- lation	7	13	15	8	20	5	9	7	3	1
3.b. Unweighted euclidean formu- lation	4	13	6	10	16	6	12	7	8	5
3.c. Weighted city-block formulation	10	9	21	12	10	9	10	3	3	2
3.d. Weighted euclidean formulation	7	10	13	10	9	11	7	9	6	4

#### Discussion and Conclusions

There are several important conceptual, methodological, and managerial implications attached to the above findings. The superiority of the ideal brand model leads one to conclude that students do form an image of the ideal school they would like to attend and evaluate existing schools on the basis of their proximity to the ideal. Furthermore, their ideal image seems rather precise since even slight deviations strongly affect preferences, as indicated by the superior performance of the exponential formulation over the euclidean and parabolic models. A possible explanation of this phenomenon involves an examination of the nature of the service

investigated in this research. The decision to enroll in a graduate business school represents a substantial financial and psychological commitment. As a result, the level of risk involved in such a decision is likely to be high as well as the need for information. For these reasons, students take the time to develop judgment criteria which lead them to a conceptualization of the ideal brand. The extent to which the results obtained in this research are generalizable to other similar types of products (of the high risk-high investment type) is unknown but, if confirmed by additional studies, these results would mean that a firm involved in the marketing of such a good would be well advised to analyze in detail the characteristics of the ideal brand, as perceived by its consumers, as well as the perceived characteristics of its own product and then design and implement an appropriate strategy for reducing the existing gap. In so doing, the firm may act at several levels of analysis including that of market segments. In this study, however, it was found (from an analysis of standard deviations attached to semantic profiles) that there was no more intersubject difference regarding the perception of the ideal brand than there was concerning existing schools. An undifferentiated strategy would therefore seem appropriate.

The second finding, that of the good relative performance of simple models, suggests that there is a rather severe limit on the amount of information that students can or wish to process before they establish their preferences. The obvious implication is that school administrators probably need not develop extensive lists of attributes unless they want to analyze specific dimensions (which they may sometimes need as a guide to brand positioning or advertising theme development, for example). What seems more important is to identify those few dimensions (actual and perceived) that really count, and carefully appraise the position of their school on each of these dimensions.

The last finding (limited impact of attribute weights) is in keeping with recent results obtained in the marketing literature (Sheth and Talarzyk, 1972 ; Moinpour and MacLahlan, 1971 ; Churchill, 1972 ; Scott and Bennett, 1971 ; Sheth, 1973). Along with Sheth and Talarzyk (1972), one may argue that the relative importance of the semantic scales is implicitly taken into account in the formulation of belief statements. In other words, a student will not assign an extreme position to a school unless he feels the scale is relatively important. Conversely, for unimportant scales, he will mainly rely upon intermediate positions. Sheth and Talarzyk have supported their augmentation by showing that the belief and weight components were substantially correlated, which violates one of the fundamental assumptions of the expectancy-value model. For practical purposes, it would therefore appear that an explicit consideration of these weights is generally superfluous.

## FOOTNOTES

1. The author expresses his thanks to Professor Charles W. Hofer of Northwestern University and Professor Susan P. Douglas of CESA for their helpful comments on an earlier draft. Financial support from the Fondation Nationale pour l'Enseignement de la Gestion des Entreprises, Paris, and the Graduate School of Management, Northwestern University is gratefully acknowledged.
2. Bernard Dubois is Assistant Professor of Marketing, Centre d'Enseignement Supérieur des Affaires (CESA), Jouy en Josas, France.
3. For an example of an empirical study which relates attitude theory to brand preference, see Bass and Talarzyk (1969).
4. A review of some of the contributions made by decision theorists to the analysis of brand preference is available in Russ (1971).
5. For a concise review of the contributions of micro-economics to the analysis of consumer behavior, see Nicosia (1966). A more detailed treatment of the underlying postulates is contained in Henderson and Quandt (1958).
6. See for example Coombs (1964).
7. For an analysis of the ideal point model, as applied to brand preferences, see Taylor (1967); for a comprehensive review of studies done on attribute and integrative models on a marketing context, see Pessemier and Wilkie (1972).
8. Two exceptions however are Russ (1971) and Bass, Pessemier, and Lehmann (1972). Yet, the range of models investigated in these studies is not as comprehensive as the one presented in this research.
9. For further details, see Dubois (1973).

## REFERENCES

- Alexis, M., Haines, G. H., Jr., and Simon, L. Consumer information processing: the case of women's clothing. In R. L. King (Ed.), Marketing and the new science of planning. Chicago: American Marketing Association Conference Proceedings, 1958.
- Bass, F. M., Pessemier, E. A., and Lehmann, D. R. An experimental study of the relationships between attitudes, brand preferences, and choice. Behavioral Science, 1972, 17, 532-41.
- Bass, F. M., and Talarzyk, W. W. A study of attitude theory and brand preference. In P. R. McDonald (Ed.), Marketing involvement in society and the economy. Chicago: American Marketing Association 1969.
- Bettman, J. R. Behavioral simulation models in marketing systems. Unpublished doctoral dissertation, Yale University, 1969.
- Churchill, G. A. Jr. Linear attitude models: A study of predictive ability. Journal of Marketing Research, 1972, 9, 423-6.



- Clarkson, G. P. E. A. A model of the trust investment process. In E. A. Feigenbaum and J. Feldman (Eds.), Computers and thought. New York: McGraw-Hill 1963.
- Comrey, A. L. A proposed method for absolute ratio scaling. Psychometrika, 1950, 15, 317-25.
- Coombs, C. H. A theory of data. New-York: Wiley and Sons 1964.
- Dubois, B. A comparative analysis of brand image and brand preference models and measurement methods. Unpublished doctoral dissertation Northwestern University, 1973.
- Einhorn, H. J., and Gonedes, H. J. An exponential discrepancy model for attitude evaluation. Behavioral Science, 1971, 16 (4), 33-9.
- Fishbein, M. A. 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 and Sons. 1967. Pp. 238-40,
- Hansen, F. Consumer choice behavior: An experimental approach. Journal of Marketing Research, 1969, 6, 436-43.
- Henderson, M. J. and Quandt, R. E. Microeconomic theory. New York: McGraw Hill. 1958.
- Lehmann, D. R. Television show preference: Application of a choice model. Journal of Marketing Research. 1971, 8, 47-55.
- Meftessel, M. A. A proposal for quantitative reporting of comparative judgments. Journal of Psychology. 1947, 24, 229-35.
- Moinpour, R., and MacLachlan, D. L. The relationships among attribute and importance components of Rosenberg-Fishbein type attitude model: An empirical investigation. In D. M. Gardner (Ed.), Proceedings of the second annual conference. College Park, Maryland: Association for Consumer Research, 1971. Pp. 365-75
- Nicosia, F. M. Consumer decision processes. Englewood Cliffs, N. J.: Prentice-Hall. 1966.
- Osgood, C., Suci, G. J., and Tannenbaum, P. H. The measurement of meaning Urbana, Ill.: University of Illinois Press, 1957.
- Pessemier, E. A., and Wilkie, W. L. Multi-attribute choice theory: A review and analysis. Paper presented at the ACR/AMA workshop on buyer/consumer information processing, University of Chicago, November, 1972.
- Rosenberg, M. J. Cognitive structure and attitudinal affect. Journal of Abnormal and Social Psychology, 1956, 53, 367-72
- Russ, F. A. Consumer evaluation of alternative product models. Paper presented at the American Marketing Association Fall Educators' Conference, Boston, 1970.

- Russ, F. A. Evaluation process models and the prediction of preference. In D. M. Gardner (Ed.), Proceedings of the second annual conference. College Park, Maryland: Association for Consumer Research, 1971. Pp. 256-61.
- Scott, J. E. and Bennett, P. D. Cognitive models of attitude structure: "Value importance" is important. In F. C. Allvine (Ed.), Combined proceedings. Chicago: American Marketing Association, 1971. Pp 256-61.
- Sheth, J. N. Brand profiles from beliefs and importances. Journal of Advertising Research, 1973, 13, 37-43.
- Sheth, J. N. and Talarzyk, W. W. Perceived instrumentality and value importance as determinants of attitudes. Journal of Marketing Research, 1972, 9, 6-9.
- Taylor, J. R. An empirical evaluation of the application of unfolding theory to market segmentation. In R. Moyer (Ed.), Changing marketing systems. Chicago: American Marketing Association, 1967. Pp. 288-96.
- Torgerson, W. S. Theory and methods of scaling. New York: Wiley and Sons, 1958.
- Wilkie, W. L., and Weinreich, R. P. Effects of the number and type of attributes included in an attitude model: More is not better. In M. Venkatesan (Ed.), Proceedings of the third annual conference. Iowa City, Iowa: Association for Consumer Research, 1972. Pp. 325-40.



SIMILARITIES AND DIFFERENCES OF GENERALIZED BRAND  
ATTITUDES, BEHAVIORAL INTENTIONS, AND  
REPORTED BEHAVIOR

Arch G. Woodside  
University of South Carolina

James D. Clokey  
Jos. Schlitz Brewing Company

and Joan M. Combes  
University of South Carolina

The first and third hypotheses of the study were supported: multi-brand/multi-attribute models more accurately classified consumer reported behavior than single brand/multi-attribute models and consumer beliefs and evaluations more accurately classified consumers' reported behavior than overall attitudes or brand intentions when complex cognitive structures existed for the consumer. The second hypothesis was not supported: normalization of belief and evaluation ratings increase correct classification of consumers' reported behavior versus the use of raw data. The theoretical and management needs for group versus individual analysis are discussed.

Attitude in consumer behavior has been defined in a multi-brand and multi-dimensional context: a cognitive state that, on a number of dimensions, reflects the extent to which the buyer prefers a brand in terms of the attributes of each brand in his evoked set in relation to other brands in the set (Howard and Sheth, 1969). However, most theoretical and operational definitions of attitude have been formulated for one brand at a time in isolation of other brands (Bhagat, Raju, and Sheth, 1974; Sheth, 1974; Sheth, 1973; Harrell and Bennett, 1974; Bass and Talarzyk, 1972; Sheth and Talarzyk, 1972). Woodside and Clokey (1974) have formulated a set of multi-brand/multi-attribute (MBMA) models of attitude with the assumption that a consumer forms an attitude toward a brand based on attributes of that brand and attributes of other brands in the same product class and in the consumer's evoked set of brands. Clokey and Woodside (1974) empirically have tested and compared predictive and diagnostic capabilities of MBMA models with previously developed single-brand/multi-attribute (SBMA) models; the predictive accuracies and explanatory powers of the MBMA models were dramatically greater than the SBMA models.

Bass and Talarzyk (1972), Bass and Wilkie (1973), and Wilkie and Pessemier (1973) have discussed potential difficulties of using cross-sectional analysis for predicting consumer attitudes because of scale level differences in belief and evaluation ratings between consumers and they recommend two methods as means to eliminate these difficulties: adjusting ratings for within-subject variance in responses (normalization) and individual-level analysis to correctly classify consumers by preferred brand. Bass and Wilkie (1973) conclude that "predictions of brand preference ranking for each respondent can be used to judge the predictive validity of cross-sectional analysis as compared to the theoretically preferred method of within-individual analysis. However, within-individual analysis is always the superior method of analysis when relative attitudes are used to predict preferences or choice, since this method does not imply respondent homogeneity, a very strong assumption necessarily implied by cross-sectional analysis."

While normalization of data for within-subject variance in rating attributes or intentions may be useful in increasing the levels of correctly classifying consumers by preferences, the theoretical preference of using within-individual instead of cross-sectional analysis is sometimes doubtful since it is important to recall "that homogenous behavior of subgroups of consumers (market segments) have always been of central concern to marketing scholars and practitioners. Therefore, finding groups that are 'homogenous enough' for cross-sectional analyses may be as important as improving the ways in which individual behavior may be more accurately modeled" (Wilkie and Pessemier, 1973). The very strong assumption implied by all types of cross-sectional analysis is theoretically and especially managerially useful for developing generalizations for subgroups of consumers, defining market segments, and planning marketing strategies, given that the cross-sectional analysis has produced substantially accurate predictions of brand preference or choice.

Sheth (1969) has demonstrated that if a complex cognitive structure exists behind the unidimensional affect-type attitude, then evaluative beliefs are better predictors of behavior as compared to attitude; when the cognitive structure behind the attitude is itself simple, both the evaluative beliefs and attitude predicted behavior to the same extent.

Comparisons of the predictive accuracies of MBMA versus SBMA models, raw versus normalized ratings, and evaluative beliefs versus unidimensional affect-type attitude (overall attitude) versus brand intention (BI) for prediction of reported behavior (RB) are examined in the present study. Specifically, the study focuses on the predictive and diagnostic merits of MBMA and SBMA models using raw and normalized ratings for consumers' beliefs that brands possess certain attributes, evaluations of these attributes, overall attitudes toward the brands, intentions to purchase the brands, and reported behavior with using the brands.

The hypotheses of the study are: (1) MBMA models more accurately classify consumer RB than SBMA models, (2) normalized ratings increase correct classification of consumers' RB, and (3) consumer beliefs and evaluations more accurately classify consumers' RB than overall attitudes or brand intentions when complex cognitive structures exist for the consumer.

#### Method

Data for testing the hypotheses were collected on a sample of 325 male household heads in the University of South Carolina's Consumer Panel. Families in the panel were selected on a quota basis of representative demographic characteristics of South Carolina.

Beliefs for five brands of beer were collected. Four of these brands were selected because of their significant market share and the fifth because of its recent entry into the market.

Beliefs and evaluations of the product category were collected for 18 attributes. The 18 attributes were selected following factor analyses of data from previous brand perception and taste studies performed by the Jos. Schlitz Brewing Company. List of potential attributes for brand perceptions were developed from previous in-depth interviewing and taste studies of small, informal, consumer groups.

Belief was operationally defined by the following statement made to the respondents and followed by an example:

I would like you to tell me how well you think each phrase describes each of the brands listed below. Please do this by putting the letter shown for each brand on one of the spaces provided. Even if you haven't tried the brand, from what you may know or think about it, indicate how well you believe the phrase describes the brand.

EXAMPLE

Not for  
Young People

For Young  
People

   :           :          C   :          D   :           :          AE   :          B   :

This would indicate that you believe Brand B is more for young people than other brands. Notice that you can place more than one brand in one space. Please take your time in answering. [Different letters were actually used in study for the specific brands studied.]

Although it seems unlikely that a consumer would actually weight 18 attributes into his decision process, it is likely that different groups of consumers may use different sets of attributes as well as different brands to determine their brand choice. Belief scales were scored from -3 to +3.

Consumer evaluations for each attribute for beer were collected by use of the following scale scored -3 to +3:

I prefer a strong beer.

Strongly Disagree       :       :       :       :       :       :       :    Strongly Agree

Overall attitude for each of the five brands of beer was collected using the following scale scored -3 to +3:

In general, I like \_\_\_\_\_ very much.

Strongly Disagree       :       :       :       :       :       :       :    Strongly Agree

Intention to purchase for each of the five brands was collected using the following scale scored -3 to +3:

I will buy some \_\_\_\_\_ during the next four weeks.

Very Unlikely           :       :       :       :       :       :       :    Very Likely

Reported behavior was collected using the following question with 10 brands presented for possible response, plus an "other" category:

Overall, considering all the brands of beer that you drink at home, in someone else's home, in restaurants, bars, or taverns, which one brand do you drink most ("X" only one box below).

A number of different forms of expectancy-value models were examined in the research study with RB as the dependent variable:

- (1)  $RB_j = \sum_{i=1}^n \alpha_{ij} B_{ij}$  (SBMA, belief only, raw data)
- (2)  $RB_j = \sum_{i=1}^n \alpha_{ij} B_{ij} N_{ij}$  (SBMA, belief only, normalized)
- (3)  $RB_j = \sum_{i=1}^n \sum_{j=1}^m \alpha_{ij} B_{ij}$  (MBMA, belief only, raw data)
- (4)  $RB_j = \sum_{i=1}^n \sum_{j=1}^m \alpha_{ij} B_{ij} N_{ij}$  (MBMA, belief only, normalized)
- (5)  $RB_j = \sum_{i=1}^n \alpha_{ij} (B_{ij} a_i)$  (SBMA, belief x evaluation, raw data)
- (6)  $RB_j = \sum_{i=1}^n \alpha_{ij} ((B_{ij} a_i) N_{ij})$  (SBMA, belief x evaluation, normalized)
- (7)  $RB_j = \sum_{i=1}^n \sum_{j=1}^m \alpha_{ij} (B_{ij} a_i)$  (MBMA, belief x evaluation, raw data)
- (8)  $RB_j = \sum_{i=1}^n \sum_{j=1}^m \alpha_{ij} ((B_{ij} a_i) N_{ij})$  (MBMA, belief x evaluation, normalized)
- (9)  $RB_j = \sum_{j=1}^m \alpha_j A_j$  (overall attitude, raw data)
- (10)  $RB_j = \sum_{j=1}^m \alpha_j A_j N_j$  (overall attitude, normalized)
- (11)  $RB_j = \sum_{j=1}^m \alpha_j BI_j$  (behavioral intention, raw data)
- (12)  $RB_j = \sum_{j=1}^m \alpha_j BI_j N_j$  (behavioral intention, normalized)

where:

$B_{ij}$  = belief  $i$  about brand  $j$ ; that is, "how well you believe the phrase describes the brand."

$a_i$  = the evaluative aspect of  $B_{ij}$ ; that is, "I prefer a \_\_\_\_\_ beer."

$A_j$  = overall attitude toward brand  $j$ ; that is, "in general, I like Brand  $K$  very much."

$BI_j$  = brand intention toward  $j$ ; that is, "I will buy some  $K$  during the next four weeks."

$RB_j$  = reported behavior toward  $j$ ; that is, "which one brand to you drink most."

$\alpha$  = coefficient estimated by multivariate analysis, e.g., multiple discriminant analysis.

$n$  = number of attributes

$m$  = number of brands

Belief only expectancy models (1 - 4) were included in the models to be tested since some empirical evidence exists that the use of evaluations or "importances" may not substantially improve the classification accuracies of brand choice (Clokey and Woodside, 1974; Sheth, 1973; Bass and Wilkie, 1973). The use of belief x evaluative models may be more intuitively appealing since consumers may not always implicitly give evaluations when judging beliefs; consequently, the attributes which are salient (large  $|\alpha_{ij}|$ ) may be different for belief versus  $B_{ij} a_i$  models.

Only disaggregated forms of the expectancy models were formulated since all previous research comparisons of disaggregate with aggregate models have shown the disaggregate to be more accurate in predicting brand choice and richer in explaining what product attributes contribute to the consumers' preference and intention to purchase (Harrell and Bennett, 1974; Bhagat, Raju, and Sheth, 1974; Sheth, 1973).

### Results

A total of 308 questionnaires were returned (95% response rate). A total of 92 (30%) of these respondents reported that they did not drink beer. A total of 112 (36%) respondents completed all sections to the questionnaire. Another 104 (34%) returned partially completed questionnaires due to lack of experience with the new brand and low usage in drinking beer. Analysis of the results is based on the 112 respondents completing all sections of the questionnaire.

Data from the study were analyzed by stepwise multiple discriminant analysis (MDA) since the dependent variable (RB) was nominal and the independent variables can be assumed to be interval. This procedure reduced the problem of covariance in the independent variables used in the models tested, i.e., the correlation coefficients between the independent variables entered in initial steps (up to 15 steps) of the MDA functions were rarely statistically significant.

A total of 9 consumer groups by brand choice were found among the 112 respondents. Correct classifications for  $RB_j$  for the 12 models tested are shown in Table 1. The SBMA percentages shown in Table 1 are averages of correct classifications for brand choice groups for Brands K, L, M, and N, the four largest brand groups. MDA was performed for SBMA models for belief data for each of these four brands, i.e., 4 MDA's were run for each SBMA model based on the  $B_{ij}$  scores for each brand and the correct classification for the respective brand groups was recorded. The correct classifications for the MBMA models are across all 9 consumer groups by brand choice. The correct classifications would be slightly higher for the MBMA models shown in Table 1 if only the 4 major consumer groups by brand choice were used to calculate correct classifications. Therefore, the results in Table 1 are biased in favor of the SBMA models.

The MBMA models substantially out perform the SBMA models by 10% or greater increases in correct classifications in all cases, e.g., MBMA for equation 7 and 10 variables had 66.7% correct classification of consumers into brand groups while SBMA for equation 5 and 10 variables had 51.3%. These results confirm the first hypothesis.



TABLE 1

Analysis of Correct Classifications of Models  
Tested for 5 and 10 Variables

Model	Equation	Percent Correctly Classified	
		5 Variables	10 Variables
SBMA, $B_{ij}$ , Raw Data	1	32.2% <sup>1</sup>	58.5% <sup>1</sup>
SBMA, $B_{ij}$ , Normalized	2	33.1 <sup>1</sup>	57.4 <sup>1</sup>
MBMA, $B_{ij}$ , Raw Data	3	43.8	71.6
MBMA, $B_{ij}$ , Normalized	4	49.1	73.3
SBMA, $B_{ij} a_i$ , Raw Data	5	32.6 <sup>1</sup>	51.3 <sup>1</sup>
SBMA, $B_{ij} a_i$ , Normalized	6	27.5 <sup>1</sup>	45.2 <sup>1</sup>
MBMA, $B_{ij} a_i$ , Raw Data	7	50.0	66.7
MBMA, $B_{ij} a_i$ , Normalized	8	45.5	69.6
Overall Attitude, Raw Data	9	50.0	n.a.
Overall Attitude, Normalized	10	35.7	n.a.
Behavioral Intention, Raw Data	11	55.4	n.a.
Behavioral Intention, Normalized	12	51.8(4 variables)	n.a.

<sup>1</sup>Average correct classifications for four brand functions.  
n.a. = not applicable.

TABLE 2

Percent Correct Classification by Brand  
Group for Five Variables

Model (Equation)	Brand Group			
	K	L	M	N
MBMA, $B_{ij} a_i$ , Raw Data (7)	59.1%	57.1%	47.6%	23.8%
Overall Attitude, Raw Data (9)	54.4	57.1	61.9	72.7
Behavioral Intention, Raw Data (11)	36.4	71.4	71.4	100.0
Sample Size	22	28	21	11

However, the discriminant analyses computed in this study used all the data to estimate the discriminant coefficients because of the relatively small sample sizes and the desire to obtain the most stable estimates of the coefficients available. This procedure produces an upward bias in the percentage correctly classified. The modified  $V_2$  validation procedure suggested by Frank, Massy, and Morrison (1965) was applied to the data. The data for each respondent were randomly assigned into the brand groups while the sizes of the groups were kept equal to their actual values. MDA was performed on these data. This procedure was performed twice for each model. The average correct classifications after random assignment were substantially below actual results (14% less or greater) in all comparisons.

The second hypothesis was not supported by the analyses. The normalized models did not substantially increase or decrease the percentages of correct classifications compared with the models using the raw data. Evidently, similar use of the possible scaled responses were used by the respondents for  $B_{ij}$ ,  $a_i$ , and BI scores. The normalized model for overall attitude actually produced a decrease in correct classification compared with the raw data model (35.7% versus 50.0% correctly classified, respectively).

The use of BI, raw data, scores produced the highest percentages of correct classification (55.4%) when 5 variables were entered into the MDA functions. To test the third hypothesis, the two consumer brand groups with the most accurate correct classifications were compared with the two brand groups with the least accurate correct classifications for Brands K, L, M, and N for the BI and overall attitude, raw score models versus the MBMA,  $B_{ij}$   $a_i$ , raw data model (equation 7). Equation 7 should produce more accurate classifications of consumers by brand choice than equations 9 or 11 for those brands classified relatively well. Results are shown in Table 2.

The third hypothesis is confirmed as shown in Table 2. Brand N drinkers was the group most accurately classified by the BI model (100%) and least accurately classified by the MBMA,  $B_{ij}$   $a_i$ , model (23.8%). Brand K was the most accurately classified by the BI model. Accuracy decreases for the MBMA,  $B_{ij}$   $a_i$ , model as accuracy increases for the BI model across the four brands. The same finding occurs for the MBMA,  $B_{ij}$   $a_i$ , model versus the overall attitude model.

The belief only ( $B_{ij}$ ) versus the belief times evaluation models ( $B_{ij}$   $a_i$ ) were somewhat more accurate in all cases when 10 variables were entered into the MDA functions. The two types of models performed similarly with 5 variables. The greater accuracy found for belief only models has occurred in previous research and suggests the need for comparisons of differences in the saliencies of variables in forming both types of disaggregate models. The multiplicative models may be more useful for diagnostic reasons even though they are somewhat out performed in accuracy by the belief only models.

Examples of the additional diagnostic capabilities of the MBMA models compared with the SBMA attribute models have been provided elsewhere (Clokey and Woodside, 1974). The data shown in Table 3 are examples of the type of information analyzed from MBMA,  $B_{ij}$   $a_i$ , models (equation 7).

Variable 5 for three different brands (1 of 18 attributes) was the most salient variable in discriminating consumer groups along with variable 18 (for Brand N) and variable 17 (for Brand L). Note that the highest mean for a particular variable and brand is for the brand drinkers of that brand, e.g.,  $\bar{X} = 6.59$  for brand K drinkers for variable 5 for brand K which is greater than any other mean for brand group K and for variable 5(K). Thus, attributes related

TABLE 3

Means ( $\bar{X}$ ), Standard Deviations (s), Standardized Betas,  
and Confusion Matrix for MBMA,  $B_{ij}$   $a_j$ , Model for  
Five Variables Entered into the MDA Functions

Variable (Brand)	Brand Group							
	K (n = 22)		L (n = 28)		M (n = 21)		N (n = 11)	
	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s	$\bar{X}$	s
5 (N)	1.45	4.95	1.86	4.56	.00	4.95	7.91	2.02
18 (N)	.05	5.04	-.21	3.53	-3.33	3.62	-1.82	6.03
5 (M)	.32	4.42	1.61	5.38	5.57	3.59	2.92	4.04
5 (K)	6.59	4.56	.11	6.09	.86	5.53	3.09	3.30
17 (L)	2.73	3.43	6.11	3.26	.24	4.69	3.73	4.63

	Standardized Betas			
	K	L	M	N
5 (N)	-.15	.69	.06	.88
18 (N)	.30	-.07	-.79	-.92
5 (M)	.01	.47	.97	.57
5 (K)	1.11	-.54	-.06	-.26
17 (L)	.38	1.16	-.64	1.12

Actual	Confusion Matrix in Percent					
	Predicted					
	K	L	M	N	Other	Total
K	59.1%	4.5%	0.0%	0.0%	36.4%	22
L	10.7	57.1	10.7	3.6	17.9	28
M	4.8	4.8	47.6	4.8	38.1	21
N	0.0	0.0	0.0	45.5	54.5	11

to the brand drunk appear to be salient for the consumer of that brand. Other attributes for other brands also appear to be salient, e.g., the standardized beta ( $b^*$ ) for variable 17 for brand L was discriminating not only for brand L drinkers but also for brand N. In fact, among the five variables used in calculating the MDA functions, 17(L) was more salient than all other variables in discriminating the brand N group.

The confusion matrix shown in Table 3 presents the correct and incorrect classifications produced by the discriminant functions. The off-diagonal percentages provide indications of degree of associations between brands, e.g., brands K, M, and N appear to be associated with other brands while brand L drinkers are somewhat likely to be classified as brand K or M drinkers. Potential brand switching behavior and reasons for such behavior may be indicated by confusion matrixes and  $b^*$ .

### Implications

Methods of answering important managerial and theoretical questions are suggested by the use of MBMA models of consumer attitudes and brand choice. What attributes of brand X are used by brand X consumers in choosing brand X? What attributes of brand Y are used by brand X consumers in choosing brand X? What attributes of brand X are used by brand Y consumers in choosing brand Y? What attributes of brand Y are used by consumers of brand Y in choosing brand Y? What brand consumer groups used attributes of brand X in choosing their brand? What brand consumer groups do not use attributes of brand X in choosing their brand? MBMA models provide an approach to defining competing brands for a particular brand based upon consumer beliefs and evaluations, as well as the attributes salient for specific brands among such consumers.

Normalization may not necessarily increase classification accuracies or the variation explained in data. However, the question of whether or not the use of raw data suppresses the significance of the results of analysis should be answered. Some consumers (e.g., light users) may tend to systematically respond to scaled items differently than others. Further research on this subject is needed.

The need for obtaining consumers' beliefs and evaluations about a brand as well as overall attitude and intention has been supported; such beliefs and evaluations may be especially relevant when complex cognitive structures exist for the consumer in his brand choice.

### FOOTNOTES

1. Associate Professor and Program Director of Marketing.
2. Senior Research Manager, Business Research Department.
3. Candidate for Doctor of Philosophy in Business Administration.

### REFERENCES

- Bass, F. M. & Wilkie, W. L. A comparative analysis of attitude predictions of brand preference. Journal of Marketing Research, 1973, 10, 262-269.

- Bass, F. M. & Talarzyk, W. W. An attitude model for the study of brand preferences. Journal of Marketing Research, 1972, 9, 93-96.
- Bahgat, R. S., Raju, P. S. & Sheth, J. N. The prediction of consumer buying intentions: a comparative study of the predictive efficacy of two attitudinal models. Paper presented at the 82nd Annual Convention of the American Psychological Association, August - September 1974.
- Clokey, J. D. & Woodside, A. G. Multi-brand/multi-attribute versus single-brand/multi-attribute attitude models: Some results. Combined Proceedings, American Marketing Association, 1974, forthcoming.
- Frank, R. E., Massy, W. F., & Morrison, D. G. Bias in multiple discriminate analysis. Journal of Marketing Research, 1965, 2, 250-258.
- Harrell, G. D. & Bennett, P. D. An evaluation of the expectancy value model of attitude measurement for physician prescribing behavior. Journal of Marketing Research, 1974, 11, 269-278.
- Howard, J. A. & Sheth, J. N. The theory of buyer behavior. New York: John Wiley & Sons, 1969.
- Sheth, J. N. Attitude as a function of evaluative beliefs. Paper presented at the American Marketing Association Consumer Workshop, Columbus, Ohio, August 1969.
- Sheth, J. N. Brand profiles from beliefs and importances. Journal of Advertising Research, 1973, 13 (1), 37-42.
- Sheth, J. N. & Talarzyk, W. W. Perceived instrumentality and value importance as determinants of attitudes. Journal of Marketing Research, 1972, 9, 462-465.
- Wilkie, W. C. & Pessemier, E. A. Issues in marketing's use of multi-attribute attitude models. Journal of Marketing Research, 1973, 10, 428-441.
- Woodside, A. G. & Clokey, J. D. Multi-attribute/multi-brand attitude model for the study of consumer brand choice. Journal of Advertising Research, 1974, 14 (5), 33-40.

## EXPLAINING BEHAVIOR WITH WEIGHTED, UNWEIGHTED AND STANDARDIZED ATTITUDE SCORES<sup>1</sup>

Sherren Waung<sup>2</sup>  
Lever Brothers Company

This investigation contrasted two alternative multi-attribute attitude models in terms of their relative ability to explain and diagnose attitude and actual behavior. The specific dependent variables were represented by physician attitudes toward prescribing specific drugs and product loyalty over eleven prescription trials. Of the two alternative models, one, an unweighted beliefs formulation, appeared to generate greater explanatory power than its weighted beliefs counterpart. This occurred in the case of explaining both attitude and product loyalty. This analysis also showed that little contribution was added by controlling for individual differences in response style via the creation of standardized scores.

### Introduction

It is apparent that the extensive development, application and refinement of multi-attribute attitude models has enabled them to assume a posture of acceptability and credibility in consumer research. Yet as it has been pointed out, major issues pertaining to conceptual development, measurement and testing remain far from resolution (Wilkie and Pessemier, 1973). This investigation attempts to shed some insight into three specific issues, namely:

- (1) Which model structures offer greater potential explanatory and diagnostic potential when product affect is the dependent variable?
- (2) Do these same structures perform satisfactorily when behavior over time represents the dependent variable?
- (3) What contribution to explanatory and diagnostic capability is made evident by the control of individual response tendencies?

### Background

The focal point of this paper is to describe the prescribing patterns of medical doctors via multi-attribute attitude models. Selection of this phenomenon for examination was generated by a number of disparate, but equally cogent concerns. One was an observation that many, indeed most marketing studies dealing with multi-attribute attitude models stopped short of relating these formulations to purchase behavior. Indeed, of the forty-two empirical studies summarized in the Wilkie and Pessemier (1973) article, only eight attempted to predict or explain actual choice decisions and even fewer did so in the context of a real world environment.

A second reason for investigating physician prescribing behavior was because it occurs in a data rich and relatively noise free environment that appears most appropriate for testing attitude behavior models. The value and extent of these characteristics are more clearly delineated if one considers how the physician product selection process might be analyzed as compared to that of the housewife as typically investigated in consumer behavior or marketing research. One basic assumption concerning consumer brand choice is that the consumer is rational (Howard and Sheth, 1969). To the extent that an individual is prone to impulse purchases or becomes satiated with a preferred brand or product, errors in predicting behavior will occur. In contrast to the housewife, decision-making by the physician is assumed to be the result of rigorous training and adherence to professional standards, both of which emphasize the formation of carefully weighed judgements. The physician being less prone to impulsive selection behavior appears to be a more suitable subject upon which to test the validity of behavioral models than the shopping housewife. The physician, when treating a common disease, often is the sole decision-maker, not likely to be directly influenced in drug selection by a professional colleague. This characteristic eliminates another potential source of error which is often difficult to control for, i.e. situations where a consumer purchase may be influenced by numerous interested individuals such as family members. The existence of a hard criterion variable, specifically recorded prescriptions, reduces sources of error that are likely to arise with purchase measures based on recall. In this regard, Parfitt (1967) has noted that attempts to recall behavior beyond the recent past produces exaggeration or oversimplification which markedly biases purchase data. Situational variables are often capable of being measured or controlled in analyzing each prescription act. Data pertaining to disease, patient age and sex, presence of drug allergies, history of previous illness and laboratory tests are available through patient records or concurrent data gathering. Numerous researchers in social psychology (Fishbein, 1967; Rokeach, 1968; Wicker, 1969) and marketing (Day, 1970; Sheth, 1970) have noted the potential of situational variables for explaining behavior not accounted for by direct attitude and multi-attribute models. Finally, since major new drug introductions are rather infrequent and most diseases (except for influenzas and respiratory infections) are not seasonal, drug markets are relatively stable. This characteristic suggests that individual prescribing patterns are likely to be more consistent over time than housewife purchases of non-durables. In addition, the high incidence of commonly treated diseases often requires physicians to repeat similar decisions within a short period of time. This short "re-prescription" cycle suggests the likelihood of an attitude change is rather small since fewer variables are likely to intervene between decisions and attitudes may have become strong and therefore resistant to change because of extensive experience with the drugs in question.

A final, but no less important basis for this research stemmed from the relative absence of key information regarding physician prescribing habits and their determinants. Attention has been drawn by the medical professional itself to the existing problems of drug selection and therapy. Dowling, a former member of the American Medical Association Council on Drugs, (1970, 1971), voiced criticism of his own medical profession for not attaining high standards in the use of therapeutic drugs. He especially deplored the reliance and dependence of his fellow physicians on commercial sources for new drug information. Two medical educators, Stolley and Lasagna (1969), emphasized the need to understand physician prescribing patterns and the rationale underlying drug selection. They attributed this need to the widespread use of new, powerful drugs, and the increasing recognition of adverse side effects. Among the questions they posed were:<sup>3</sup>

- (1) What are the factors leading to new drug adoption or persistent use of an older drug?
  - (2) Why are some drugs prescribed with the hope of producing a pharmacological effect they do not possess?
  - (3) Why are potent antibiotics still prescribed by a significant number of physicians for trivial infections?
- and
- (4) Why do certain drugs retain their popularity while most expert opinion decries their use?

#### The Theory and Hypotheses

Much controversy has concerned the alternative structures of the multi-attribute attitude models. Both the original Rosenberg (1956) and Fishbein (1963) paradigms from social psychology have spawned new reformulations, yet a common thread still remains among the initial models and their subsequent variations: to explain consumer predispositions toward objects (brands/products) via an appropriate weighting of decision choice-related beliefs. The formulations below represent the two main structures at issue, and have undergone some translation in the context of physician drug-related attitudes and behavior.

$$\text{Model I } A_k = f (b_1 [P.I.1_k \times V.I.1] + \dots + b_n [P.I.n_k \times V.I.n])$$

where:  $A_k$  = attitude toward prescribing the  $k^{\text{th}}$  drug  
 $P.I.j_k$  = the extent that the  $k^{\text{th}}$  drug meets the  $j^{\text{th}}$  selection criterion.  
 $V.I.j$  = the value importance associated with meeting the  $j^{\text{th}}$  selection criterion, i.e., the desirability of meeting that criterion.

$$\text{Model II } A_k = f (b_1 [P.I.1_k] + \dots + b_n [P.I.n_k])$$

A number of observations require mention. Both models incorporate the perceived instrumentality component of Rosenberg (1956), however, the variable shown is probably most closely related to Sheth's (1969) definition of evaluative beliefs. In this case, the term expresses the extent to which drug selection criteria (secondary goals) are met by the drug in question. Model I possesses the value importance term of Rosenberg (1956), however restated in terms of desirability rather than personal satisfaction. As the dependent variable, attitude relates to the act of prescribing a drug and therefore is similar to Fishbein's (1967) construct of attitude towards an act (Aact).

The advisability of including the value importance component is the issue at hand. In assessing the empirical support for and against inclusion of this term, Bass and Wilkie (1973) note that the bulk of evidence suggests importance weights are not likely to improve the explanatory capability of multi-attribute models, or decrease it as first proposed by Sheth and Talarzyk (1972). A number of recent studies, however, have indicated significant increases in predictive and diagnostic power attributable to the use of normalized scores (Bass and Wilkie, 1973; Ginter, 1973; Wilkie, McCann, and Reibstein, 1973).



Given this background, the research objective is to re-examine this issue by comparing the explanatory and diagnostic power of Models I and II. The two relevant null hypotheses are:

- H1 No differences exist among the proposed models in explaining physician attitudes toward prescribing the kth drug.
- H2 No differences exist among the proposed models in explaining physician product loyalty towards the kth drug.

#### Method

##### Subjects

Fifty-five physicians from four university health centers and a local community volunteered to participate in this study.

##### Procedure

Self-administered questions related to the behavioral phenomenon of interest, the treatment of a specific common disease, were posed in the context of a typical prescribing situation. Specifically, the physician was asked to imagine treating a first episode case which was to be treated prior to the positive identification of the organism and its susceptibility to different drugs. Responses to the following were elicited in the framework of this prescribing situation.

- (1) Affect (Direct). Respondents indicated the strength of their preference for prescribing three different drugs on 7-point scales. This measure was intended to reflect predisposition toward the act of using a drug in a well-defined situation, and not toward the drug itself in a general context.
- (2) Value Importance. Respondents indicated how desirable they consider each of seven choice criteria in selecting a drug under the specific situational conditions defined.
- (3) Perceived Instrumentality. Physicians were asked to indicate the degree to which each drug meets each of the drug choice criteria on a 7-point scale. Five attributes, evolved from exploratory research and pre-tests, were used.

Patient medical records were screened until information pertaining to eleven prescribing occasions or trials were gathered. From these data two measures of product loyalty were formed, one simply being the proportion of total prescriptions devoted to each drug. The other measure was derived by calculating loyalty scores based on a factor analytic procedure proposed by Sheth (1970). This latter measure reflected, for each physician, the pattern and frequency followed in prescribing each drug.

#### Results

##### Intrasubject Reliability

Twelve perceived instrumentality judgements were repeated after completion of the main questionnaire. The mean reliability was .91, the range .38 to 1.0. A final sample size of forty-five was determined as a result of these specific respondent assessments.

## Validation Assessment of Affect Measure

A multitrait - multimethod approach (Campbell and Fiske, 1959) was employed to assess the validity of affect. The mean  $r$  of affect with product loyalty was .55 ( $p < .01$ ) thus adding assurance that the measure was appropriate.

## Testing of Hypothesis 1

Prior to testing Hypothesis 1, two transformations were performed on the data in order to account for possible individual differences in response style. These within subject standardizations of raw responses were suggested by a number of considerations. Osgood, Suci and Tannenbaum (1957) in their research on the semantic differential noted numerous instances where response style differences among individuals were apparent. While discussing the benefits and pitfalls of within subject standardization, they noted that better educated subjects tend to use the intermediate positions relatively more frequently than the polar or neutral positions (Stagner and Osgood, 1946). In addition, intelligence scores appear to be related to position usage in responding to the semantic differential, subjects with lower scores being more polarized in their judgements (Kerrick, 1954). Heise (1969) in a recent review of methodological issues in semantic differential research noted the presence of important differences in response styles, particularly a tendency for some subjects to use end points more often than the discriminatory intermediate positions. Peabody (1962) in turn concluded that this propensity is a stable trait of individuals over time and over different sets of concepts. Finally, Bass and Wilkie (1973) supported hypotheses that cross-sectional regression models utilizing normalized scores would yield higher proportions of explained variance and more significant coefficients than those employing raw responses. The authors attributed their findings to the fact that: (1) preference ranks are normalized by individual, so that the distribution of the dependent variable will differ for the separate brand regressions; (2) scale differences in belief and importance ratings may not represent true scores. These two conditions result in regressions which attempt to explain brand preference (ranked) via belief and importance scores, which reflect individual response style differences.

In view of the above mentioned considerations, within subject standardizations were conducted on perceived instrumentality ratings: (1) across all three products and all five attributes so as to control for differences in response style which might be reflected across all drug criteria ratings; (2) across all three products, but within each drug criteria in order to control for response style variations in perceived instrumentality ratings of products against a given drug criteria. The option of standardizing value importance (desirability) scores within each individual was given serious consideration; however, the decision to proceed with this was made contingent upon the results stemming from the use of standardized perceived instrumentality scores. This seemed consistent with the Bass and Wilkie (1973) results which demonstrated that normalization of belief, rather than importance scores, yielded substantially greater explanatory power over raw responses.

The testing of Hypothesis 1 therefore entailed formulating the two different attitude models in terms of raw scores and two different sets of standardized scores. Then a total of six regressions of affect against different attitude formulations were performed for each drug product. Results of the 18 runs regressing affect onto the proposed models are summarized on the next page.

TABLE 1  
Percent Variance Explained in Affect  
(Adjusted R<sup>2</sup>)

Predictors	DRUGS	<u>T</u>	<u>S</u>	<u>N</u>	<u>Mean R<sup>2</sup></u>
<u>Model I (Weighted PI)</u>					
	S	.49c	.19a	.14	.27
	S <sub>1</sub>	.12	.18a	.18a	.16
	S <sub>2</sub>	.19a	.06	.10	.12
	Mean R <sup>2</sup>	.27	.14	.14	.18
<u>Model II (Unweighted PI)</u>					
	S	.64c	.30b	.32b	.42
	S <sub>1</sub>	.43c	.36c	.34c	.38
	S <sub>2</sub>	.34c	.37c	.26b	.32
	Mean R <sup>2</sup>	.47	.34	.31	.37

- S - Perceived instrumentality scores in raw form.  
 S<sub>1</sub>- Perceived instrumentality scores standardized within individual across products and selection criteria.  
 S<sub>2</sub>- Perceived instrumentality scores standardized within individual across all products within each selection criteria.  
 a -  $p < .05$ , one-tail F test  
 b -  $p < .01$ , one-tail F test  
 c -  $p < .001$ , one-tail F test

The disaggregated versions provided the greatest explanatory power in physician affect toward the three drug products. The mean percent of variance explained was 37% and 18%, for Models II and I respectively. Each equation of Model II was significant with respect to explaining affect toward each drug. Specifically, 9 out of 9 Model II regression equations were significant. In contrast, Model I yielded 5 significant equations and had no single formulation which yielded a significant amount of explained variance in attitude towards each of the three products. A different assessment of relative predictive power involved F-tests for significant differences in R<sup>2</sup>. On no occasion did the weighted model generate a statistically higher R<sup>2</sup> than its unweighted counterpart. Instead, in 4 out of the 9 contrasts, the unweighted formulation was superior. No significant differences in the mean number of coefficients were found among Models I and II. Thus each was generally equivalent in terms of this operational definition of diagnostic power. To summarize, differences in ability to explain affect among the proposed models are apparent, therefore the null hypothesis is rejected. These differences were manifest in terms of percent variance explained in affect and the number of products towards which physician affect could be explained. No differences, however, were found with respect to the number of significant coefficients provided by the models.

## Testing of Hypothesis 2

Eighteen regressions of physician product loyalty (represented by respondent factor scores) onto the hypothesized attitude models were performed.

TABLE 2

Percent Variance Explained in Product Loyalty

(Adjusted  $R^2$ )

DRUGS Predictors	T	S	N	Mean $R^2$
<u>Model I (Weighted PI)</u>				
S	.30b	.19a	.12	.20
S <sub>1</sub>	.10	.08	.18a	.12
S <sub>2</sub>	.06	.04	.31b	.14
Mean $R^2$	.15	.10	.20	.15
<u>Model II (Unweighted PI)</u>				
S	.26b	.20a	.00	.15
S <sub>1</sub>	.27b	.28b	.15	.23
S <sub>2</sub>	.30b	.17a	.20a	.22
Mean $R^2$	.28	.22	.12	.20

- S - Perceived instrumentality scores in raw form.  
 S<sub>1</sub> - Perceived instrumentality scores standardized within individual across products and selection criteria.  
 S<sub>2</sub> - Perceived instrumentality scores standardized within individual across products within each selection criteria.  
 a -  $p < .05$ , one-tail F test  
 b -  $p < .01$ , one-tail F test  
 c -  $p < .001$ , one-tail F test

In contrast to previous findings, the product loyalty regressions were characterized by generally low explanatory power. For example, the mean  $R^2$  was .42 and .39 for the two equations which provided the greatest explanation in affect. In contrast, the highest  $R^2$  attained among the product loyalty regressions was .23. Seven equations of Model II proved significant, while four formulations of Model I attained significance. In terms of significantly different  $R^2$ s, once again the unweighted model held a marginal edge over the weighted version.

TABLE 3

Number of Significantly Higher  $R^2$ s Among Loyalty Regressions

Model I over Model II	0
Model II over Model I	2
Total Number of Comparisons	2

The differences in the number of significant coefficients generated were insignificant across models, once again demonstrating equal diagnostic capability via this narrow definition.

Because differences among the models were apparent with respect to their capability to explain loyalty the null hypothesis is rejected. As has been previously found, the disaggregated unweighted beliefs model provided the greater explanation in loyalty.

#### Comparison of Raw and Standard Score Regressions

A series of F-Tests were performed to determine if the  $R^2$ s produced by the raw score equations were significantly higher than those generated by the two standardized score formulations.

TABLE 4

Number of Significantly Higher  $R^2$ s

	Attitude Regressions		Loyalty Regressions	
	Model I	Model II	Model I	Model II
S over $S_1$	1	1	1	0
S over $S_2$	1	1	1	0
$S_1$ over $S_2$	0	0	0	0
$S_2$ over $S_1$	0	0	0	0
Number of Comparisons	6	6	6	6

In no instance did any standard score regression produce a higher  $R^2$  than its raw score counterpart. The obverse held true in four comparisons.

#### Stability of Results

The limited sample size of the study necessitated a comprehensive investigation of the stability of the regression results. In particular, a procedure suggested by Mosteller and Tukey (1968) was employed for the purpose of providing refined parameter estimates which incorporated direct assessments of the variability in the data. These authors contend that major sources of variation must be assessed in deriving results, citing group or cluster variation as being a major determinant of stability, particularly in the case of small samples. The jackknife is a means which allows one to determine directly the variability in a data set by estimating the same statistic computed on several overlapping groups of data. Thus, this method was employed to assess the effect of physician sample clusters upon the aggregate regression results. This was conducted by successively eliminating each physician cluster from the total sample over a series of regression analyses. As a result of this procedure,  $R^2$ s, regression coefficients and standard errors of coefficients were re-estimated for the affect and loyalty regressions. As it turned out, all re-estimated parameters were highly consistent with the original ones. In no instance were the conclusions derived from testing either of the hypotheses altered. In sum, the small sample based estimates were not altered by within group variation thereby indicating stability of the results.

### Summary and Implications

The findings contribute further support to the contention that multi-attribute attitude models can explain behavior, in this particular instance prescribing patterns over time. While the amount of variance explained in loyalty would likely shrink in a cross validation of these results, nonetheless they appear satisfactory. This appears so given the  $R^2$  of .10 typically found in social psychology studies which attempted to explain or predict behavior from attitude models (Wicker, 1969).

The control of individual differences through standardization, with a few exceptions, did not markedly enhance the explanatory power of the models. This does not mean or imply that other transformations might not have improved the results. Clearly, normalization has improved results as demonstrated in a number of recent studies. A relevant point to be raised regarding transformations, however, is that as consumer researchers experiment with such data manipulation, they should continue to report all particular circumstances where model performance is or is not improved. The necessity for this stems from the recognition that sub-populations differ in their response tendencies and that whole-scale alteration of data sets may badly distort information from these key entities.

Any claims of superiority for the unweighted over the weighted model must of course be tempered by recognizing the sample size employed in this study. Nonetheless, the pattern of greater explanatory power was consistent in explaining both affect and loyalty. This type of stability ought to serve as an evaluative criteria for assessing multi-attribute models in addition to predictive and diagnostic considerations.

### FOOTNOTES

1. The author wishes to gratefully acknowledge financial support from the **Lever Brothers Company**.
2. Group Research Manager, Lever Brothers Company, and Doctoral Candidate at the University of Illinois.
3. P.D. Stolley and L. Lasagna, Prescribing patterns of physicians. Journal of Chronic Diseases, 22, 1969, p.2.

### REFERENCES

- Bass, F.M. & Wilkie, W.L. A comparative analysis of attitudinal predictors of brand preference. Journal of Marketing Research, 1973, 10, 262-269.
- Bopp, J. A quantitative semantic analysis of word association in schizophrenia. Unpublished doctoral dissertation, University of Illinois, 1955.
- Campbell, D.T. & Fiske, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.
- Day, G.S. Buyer attitudes and brand choice behavior. New York: Free Press, 1970.
- Dowling, H.F. Medicines for man. New York: Alfred A. Knopf, Inc., 1970.

- Dowling, H.F. The prescribed environment. Saturday Review of Literature, 1971, 58-60.
- Fishbein, M. An investigation of the relationships between beliefs about an object and the attitude toward that object. Human Relations, 1963, 16, 223-239.
- Ginter, J.L. The effects of normalization on the multi-attribute attitude model. Paper presented at the Fourth Annual Meeting of the Association For Consumer Research, Boston, Mass., November, 1973.
- Heise, D.R. Some methodological issues in semantic differential research. Psychological Bulletin, 1969, 72, 6, 406-422.
- Howard, J.A. & Sheth, J.N. The theory of buyer behavior. New York: Wiley, 1969.
- Mosteller, F. & Tukey, J.W. Data analysis, including statistics. In G. Lindzey & E. Aronson (Eds.), Handbook of social psychology. (2nd ed.) Vol. 2. Reading, Mass.: Addison - Wesley, 1968.
- Kerrick, J.S. The effects of intelligence and manifest anxiety on attitude change through communications. Unpublished doctoral dissertation, University of Illinois, 1954.
- Osgood, C.E., Suci, G.J. & Tannenbaum, P.H. The measurement of meaning. Urbana: University of Illinois Press, 1957.
- Parfitt, J.H. A comparison of purchase recall with diary panel records. Journal of Advertising Research, 1967, 7, 16-31.
- Peabody, D. Two components in bi-polar scales: Direction and extremeness. Psychological Review, 1962, 69, 65-73.
- Rokeach, M. Beliefs, attitudes and values: A theory of organization and change. San Francisco: Jossey-Bass, 1968.
- Rosenberg, M.H. Cognitive structure and attitudinal affect. Journal of Abnormal and Social Psychology, 1956, 53, 367-372
- Sheth, J.N. A factor analytic model of brand loyalty. Journal of Marketing Research, 1968, 5, 395-404.
- Sheth, J.N. Attitude as a function of evaluative beliefs. Paper presented at the American Marketing Association Conference Workshop, Columbus, Ohio, 1969.
- Sheth, J.H. & Talarzyk, W.W. Relative Contribution of perceived instrumentality and value importance components in determining attitudes. University of Illinois, College of Commerce and Business Administration. Working paper number 15, 1971.
- Stagner, R., and Osgood, C.E. Impact of war on a nationalistic frame of reference: Changes in general approval and qualitative patterning of certain stereotypes. Journal of Social Psychology, 1946, 24, 187-215.
- Stolley, P.D. & Lasagna, L. Prescribing patterns of physicians. Journal of Chronic Diseases, 1969, 22, 395-405.

Wicker, A.W. Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. Journal of Social Issues, 1969, 25, 41-78.

Wilkie, W.L., McCann, J.M. & Reibstein, D.J. Halo effects in brand belief measurement: Implications for attitude model development. Paper presented at the Fourth Annual Meeting of the Association For Consumer Research, Boston, Mass., November, 1973.

Wilkie, W. L. & Pessemier, E.A. Issues in marketing's use of multi-attribute attitude models. Journal of Marketing Research, 1973, 10, 428-441.





A MULTIMETHOD APPROACH TO VALIDATING  
MULTI-ATTRIBUTE ATTITUDE MODELS<sup>1</sup>

James R. Bettman, Noel Capon and Richard J. Lutz<sup>2</sup>  
University of California, Los Angeles

The strengths and weaknesses of correlational and analysis of variance (ANOVA) approaches to validating multi-attribute attitude models are discussed. The approaches are then applied to the Fishbein and adequacy-importance models, and an adequacy-evaluation model developed from earlier findings. The adequacy-evaluation model is supported reasonably well. The Fishbein model is poorly supported by correlational approaches, and the adequacy-importance model receives little confirmation from ANOVA findings. Multiple method validation approaches are emphasized.

In recent years consumer researchers have focused considerable attention on multi-attribute models of attitude structure. Drawing from the original theoretical formulations of Rosenberg (1956) and Fishbein (1963) a variety of models have been developed, operationalized and tested (Wilkie and Pessemier, 1973). These models have mostly taken the form:

$$A_j = \sum_{i=1}^n X_{ij} Y_i \quad (1)$$

where, for an individual respondent  $A_j$  = Attitude toward brand  $j$ ;  $X_{ij}$  = Belief as to the extent that brand  $j$  possesses attribute  $i$ ;  $Y_i$  = Some evaluative aspect of attribute  $i$ , such as its goodness or badness, or its importance.

The models have varied both in their underlying theoretical formulations and in their operationalizations; however, the methods used for testing the models have been remarkably similar. Typically, a single product class is selected, and respondents are presented with sets of scales designed to measure  $Y_i$  for selected attributes and  $X_{ij}$  for chosen brands on each attribute. Finally, a global  $A_j$  measure for each brand is obtained and a cross-sectional correlation analysis performed between  $A_j$  and the index  $\sum X_{ij} Y_i$ . The percentage of explained variance has been the criterion employed to compare models, the model with highest  $r^2$  being proclaimed as the superior model. An example of this multiple model approach is found in Mazis and Klippel (1973), who compared the Fishbein, Rosenberg, and adequacy-importance models.

Bass & Wilkie (1973) have pointed out several problems with the cross-sectional correlation procedure, particularly the strong assumption of comparability of individual utilities; accordingly, they have suggested a normalization procedure for adjusting for within subject variance in cross-sectional analysis. Although  $r^2$  is increased when this procedure is used, this greater degree of fit occurs irrespective of the model to which the data are fitted. In addition, by the very form of the cross-sectional analysis, all subjects are still assumed to be employing the same model.

These difficulties with cross-sectional analysis are compelling reasons for moving to individual level analysis in examining multi-attribute attitude models. The ultimate goal is to "validate" multi-attribute attitude models by identifying which model is most appropriate as a representation of reality in a given set of circumstances. The measure of appropriateness, however, is a function of the researcher's objectives, and these objectives guide the methodology which is used in the validation procedure (Slovic and Lichtenstein, 1971). If prediction is the aim, then a regression approach (Brunswik, 1952, 1956; Hoffman, 1960) is most appropriate. On the other hand, if an understanding of the individual judgment process is the objective, then an analysis of variance procedure relying on factorial designs is most appropriate (Anderson, 1973, 1974).

The purpose of this paper is to examine two individual level approaches to validation of multi-attribute attitude models and to provide empirical data for validating three alternate multi-attribute models. Results of the two individual level approaches will be compared with the results of the standard cross-sectional analysis. This multimethod approach is a crucial step in examining model validity (Wright, 1973).

#### Validation by Regression

The individual level regression approach to validation relies on a data collection procedure similar to that outlined above for the cross-sectional analysis. However, the interpersonal utility issue is sidestepped by performing individual level regressions, using the data for each brand as an observation. Validity of a model is again determined by a correlational measure: the higher the coefficient, the more "valid" the model. This approach is in the Brunswik tradition (Brunswik, 1952, 1956) in which a major focus is the realism of the task environment. This realism is achieved by presenting scales which are meaningful to respondents in the sense that only these attributes and brands that are familiar to them are included in the questionnaire. Furthermore, by demanding that respondents perform the encoding function for both the predictor and criterion variables, the task is made similar to the judgment task used in normal purchase decisions.

However, this approach has a number of drawbacks (Anderson, 1962; Sidowski & Anderson, 1967), the most serious of which is the use of  $r^2$  as a measure of fit. Birnbaum (1973) has shown how errorless interval data conforming to a perfect linear pattern can lead to a correlation coefficient for an incorrect multiplying model greater than that for the correct adding model. Conversely, errorless interval data conforming to a perfect multiplicative pattern can lead to a correlation coefficient for an incorrect adding model greater than for the correct multiplying model. Birnbaum further indicates that errors in measurement can greatly exaggerate this effect, with the result that data fit to incorrect models could result in much higher correlations than data fit to correct models.

Another problem is that interval scaling must be assumed for  $X_{ij}$ ,  $Y_i$ , and the criterion  $A_j$ . This is a particularly crucial assumption to make for data used to test a multiplicative model; in fact,  $X_{ij}$  and  $Y_i$  should really be ratio-scaled for invariant results. Any relaxation from interval to ordinal data for the predictor variables can play havoc with degree of fit for a multiplicative model. Further, the researcher must decide

whether to code the predictor scales as unipolar (e.g., 1 to 7) or bipolar (e.g., -3 to +3), which is not necessary for the ANOVA approach.

Finally, the use of a correlational approach, even at the individual level, means that all respondents' data are being compared with a single, monolithic, model of attitude formation. That is, the model becomes the focal point, with each individual's data being evaluated on how well they "fit" the model. Because each respondent's data are handled in a pre-specified manner, there is no opportunity for the discovery of new relationships through the invaluable scientific processes of induction and serendipity. A respondent either fits the model, or he does not. While those who fit the model are valuable in providing support for the model, the data of respondents who do not fit the model are not being fully utilized. The ANOVA approach described in the next section overcomes this serious drawback.

#### Validation by ANOVA

The ANOVA approach, a methodology directed toward the analysis of individual judgment processes, is quite different from the regression approach, where the goal is prediction. The major use of ANOVA has been in studies of clinical judgment (Hoffman, Slovic and Rorer, 1968; Anderson, 1969, 1972; Goldberg, 1968) and information integration (Anderson, 1962, 1973, 1974). The methodology involves the development of a series of cue dimensions and stimulus levels on each cue dimension. These cue dimensions become orthogonal factors in a factorial design where the stimulus levels are the levels on each factor. Subjects are presented with profiles representing combinations of stimulus levels from each cue dimension and requested to make a judgment. In total they are presented with all possible combinations of the full factorial design and at least two replications. In one clinical judgment study, for instance, radiologists were presented with a  $3 \times 2^5$  complete factorial arrangement of six dimensions of information about an ulcer and asked to judge whether the ulcer was benign or malignant (Hoffman, Slovic and Rorer, 1968). They were asked for two replications of the 96 cases, 192 judgments in total.

The major advantage of this procedure is that specific tests can be developed for various types of rules for combining the dimensions: adding, averaging, multiplying, or more complex models (e.g.  $a(b+c)$  or  $ab+cd$ ) (Anderson, 1973). Both ANOVA and graphical interpretation of data are useful. The specific tests and techniques used are detailed in later sections. Further, the assumption of interval data is required only for the criterion variable, a major advantage over the regression framework. In fact, by use of functional measurement, scale values can be developed for the predictor variables rather than assumed, as in the regression methodology (Anderson, 1970).

The major disadvantages of the methodology are the lack of realism involved in the factorial task and the fatigue associated with making a large number of judgments.<sup>3</sup> The factorial data structure departs substantially from the Brunswikian notion of having cues representative of the subject's normal environment. Also, the systematic factorial structure may bias subjects toward more systematic combination rules. Finally, the subject no longer has to perform an encoding task for the predictor variables, as under the regression methodology, and the stimuli which are

developed to test multi-attribute models are devoid of content and low in mundane reality. A final disadvantage of the ANOVA approach lies in the interpretation of interactions, which may arise from problems with the response scales, such as floor and ceiling effects, as well as representing a true relationship (Anderson, 1968; Bogartz and Wackwitz, 1971).

#### Validation of Multi-Attribute Attitude Models at the Individual Level

##### Regression

As pointed out above, the data collection procedure for individual level regression analysis is similar to that for cross-sectional analysis. The researcher chooses a product class, identifies brands within that class, and generates a series of relevant attributes. Respondents are asked to rate each attribute on some evaluative dimension and each brand as to the extent to which it possesses each attribute. Finally, the respondent provides a global affect measure for each brand. The analytic procedure requires regressing the criterion variable of global affect against the sum of the products of the predictor variables, with one observation for each of the several brands for each individual. The success of the model for each individual is determined by the size of the correlation coefficient.

##### ANOVA

The ANOVA approach, described in detail elsewhere (Bettman, Capon, and Lutz, 1974), is quite different from the correlation approach. As discussed earlier, the methodology is appropriate for examining the individual judgment process when subjects are provided with sets of data, each set being appropriate for the model to be tested. The methodology is described below in terms of an example of comparing competing multi-attribute models of attitude, the Fishbein model, and the adequacy-importance model (Cohen, Fishbein and Ahtola, 1972).

The Fishbein Model may be formulated as:

$$A_j = \sum_{i=1}^n b_{ij} a_i \quad (2)$$

where  $A_j$  = attitude toward brand  $j$ ;  $b_{ij}$  = the likelihood that brand  $j$  possesses attribute  $i$ ;  $a_i$  = evaluative aspect of attribute  $i$ , its goodness or badness; and  $n$  = number of attributes.

In a similar manner, the adequacy-importance model may be expressed as:

$$A_j = \sum_{i=1}^n B_{ij} I_i \quad (3)$$

where  $A_j$  = attitude toward brand  $j$ ;  $B_{ij}$  = belief as to the extent to which attribute  $i$  is offered by brand  $j$ ;  $I_i$  = importance weight given attribute  $i$ ; and  $n$  = number of attributes.

Each model makes a number of assumptions about how subjects treat data. These assumptions, or model composition rules, as expressed by the basic algebraic formulation of the models, are as follows:

- 1) The contribution to affect for each attribute is multiplicative for each model; e.g.,  $b_{ij}a_i$  and  $B_{ij}I_i$ .
- 2) The powers of the terms in the product for each attribute are assumed to be 1 for both models.<sup>4</sup>
- 3) The product terms,  $b_{ij}a_i$  and  $B_{ij}I_i$ , add over attributes to form an overall affect measure in both models.<sup>5</sup>
- 4) The components themselves,  $a_i$ ,  $b_{ij}$ ,  $I_i$ , and  $B_{ij}$ , are coded in particular ways as part of the measurement process. Researchers using the Fishbein model have measured  $a_i$  and  $b_{ij}$  on bipolar (e.g., -2 to +2) scales, while adequacy-importance model researchers have typically used nonnegative scales (e.g., 1 to 5).<sup>6</sup>

In this paper the additive assumption of the models is ignored and it is assumed that the powers equal unity. Attention is focused on the multiplicative and coding assumptions. To test these assumptions via the ANOVA approach, the researcher could present profiles representing a hypothetical brand and ask subjects to make global affect judgments. In testing the Fishbein model, profiles of the following type could be employed:

You believe that Brand X is very likely: \_\_\_:\_\_\_:X:\_\_\_:\_\_\_:very unlikely to possess a quality which you personally feel is:  
 very good :\_\_\_:\_\_\_:\_\_\_:\_\_\_:X: very bad. To you, using Brand X would be:  
 very favorable :\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_: very unfavorable.

In testing the adequacy-importance model, profiles of the following form would be appropriate:

You believe that Brand X is very high :\_\_\_:\_\_\_:X:\_\_\_:\_\_\_: very low in possessing a quality which you personally feel is:  
 very important :\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:X: not important at all. To you, using Brand X would be:  
 very favorable :\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_:\_\_\_: very unfavorable.

Since five-point scales are used to present the profiles, there would be a total of twenty-five data profiles in the full (5x5) factorial design. Each subject would be given the task corresponding to one of these two models, and the data generated from these factorial designs examined both graphically and by ANOVA. Depending on the particular (implicit) coding scheme used by the subject and the data combination rules assumed, three simple data patterns could emerge, as shown in Figure 1.

The pattern of data in Figure 1(a) -- a set of parallel straight lines -- would be found if subjects add the component bits of data, e.g.,  $(a_i + b_{ij})$ , irrespective of how they code it. In the ANOVA there would be strong main effects but no interaction. Neither the adequacy-importance model nor the Fishbein model predicts this pattern of results.

The data pattern of Figure 1(b), a set of crossing straight lines, would be found when subjects act as though coding data in a bipolar manner and then multiplying the component bits of data (bipolar multiplying). In the ANOVA there would be no significant main effects but a strong interaction effect. Such a data pattern is predicted by the assumptions of the Fishbein model.

The data pattern of Figure 1(c) -- a diverging fan of straight lines -- would be found when subjects act as though coding data in a unipolar manner

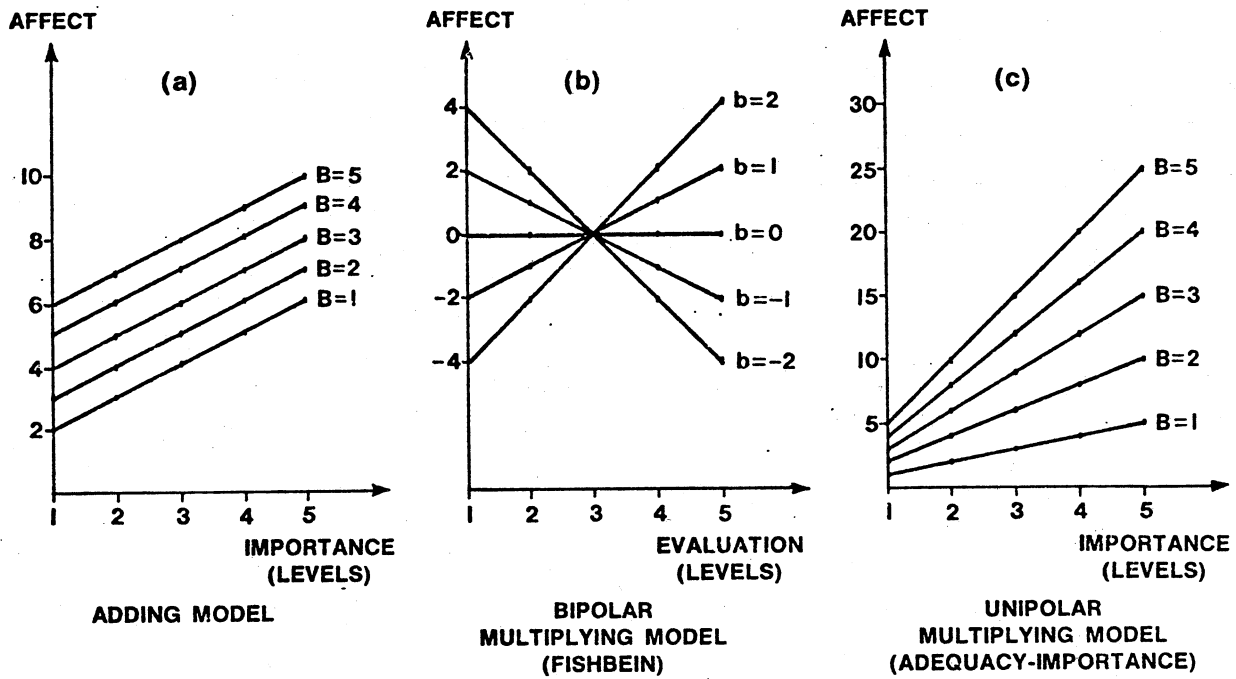


Figure 1. Theoretical Data Structures.

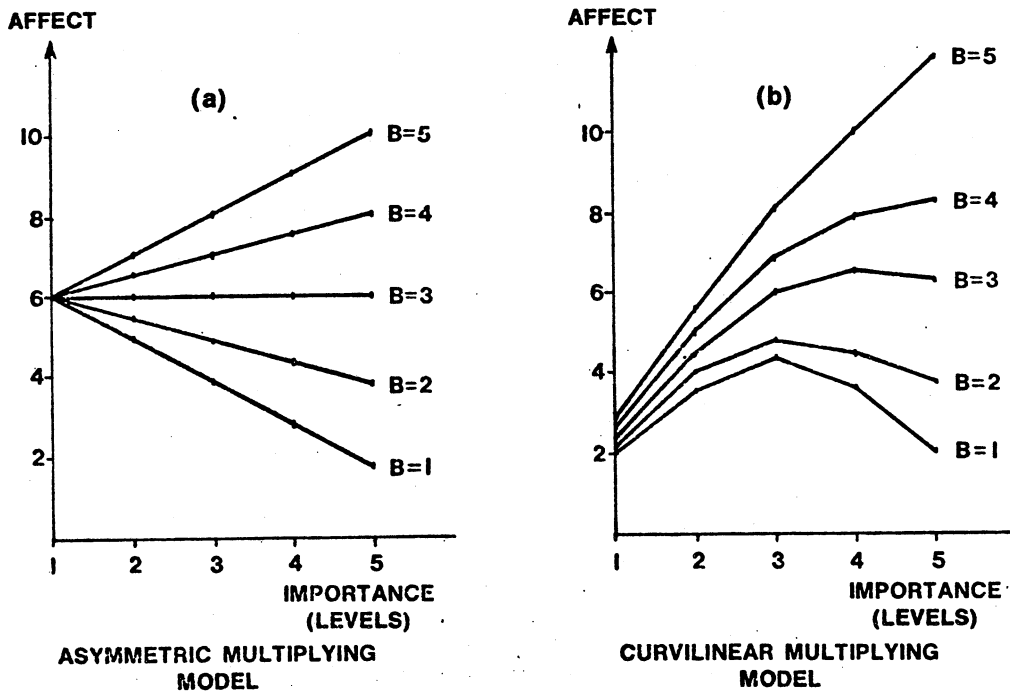


Figure 2. Additional (Non-Predicted) Data Structures.

and then multiply the component bits of data (unipolar multiplying). In the ANOVA there are both significant main effects and a significant interaction, concentrated in the bilinear component (Anderson, 1970, 1973). Such a data pattern is predicted by the assumptions of the adequacy-importance model.

#### Previous Application of the ANOVA Approach

In a previous study (Bettman, Capon, and Lutz, 1974) profiles of both the Fishbein type and the adequacy-importance type were presented to two groups of undergraduate psychology students. Following a warm-up task each subject rated 50 profiles, two replications each of the 5x5 factorial design. The data was analyzed for each individual subject as a two-factor ANOVA. In addition to the data patterns expected from the three simple models of adding, unipolar multiplying and bipolar multiplying, a number of other patterns were observed, making seven in total. Subjects were classified into these categories on the basis of both significance levels and explained variance, as measured by the  $\hat{\omega}^2$  statistic (Hays, 1963; Green, 1973). The categorization scheme is briefly outlined below. For more detail see Bettman, Capon, and Lutz (1974).

#### Category 1 - No Significant Effects - "No Effects"

Subjects falling into this category have no significant effects for either the treatments or the interaction.

#### Category 2 - Significant Main Effects - "Adding Model"

Subjects in this category exhibit at least one significant main effect, but have no significant interaction effects. Their data plots correspond to Figure 1(a) -- i.e., a set of parallel lines. The underlying theoretical model implied is addition, rather than multiplication, of the components of the model. This is in contrast to both the Fishbein and adequacy-importance models.

#### Category 3 - Significant Main Effects, Significant Interaction Concentrated in the Bilinear Term - "Unipolar Multiplying Model"

Subjects in this category exhibit two significant main effects and a significant interaction concentrated in the linear x linear (bilinear) component of the interaction (Anderson, 1970).

The data correspond to the plot in Figure 1(c) -- i.e., a diverging fan of straight lines in the upper right quadrant. The underlying theoretical model is multiplication of components coded in a unipolar fashion, which corresponds to the assumptions of the adequacy-importance model.

#### Category 4 - Major Portion of Explained Variance in Interaction - "Bipolar Multiplying Model"

The criteria which had to be met for inclusion in this category were as follows: the amount of explained variance in the interaction, as measured by  $\hat{\omega}^2$ , had to be at least 35%, and in addition, had to be at least three times greater than the explained variance for each main effect taken separately. The underlying theoretical model is multiplication of components coded in a bipolar fashion, which corresponds to the assumptions of the Fishbein model.



Category 5 - Major Portion of Explained Variance in One Main Effect and the Interaction - "Asymmetric Multiplying Model"

Subjects in this category have one strong main effect which explains roughly the same proportion of the variance as does the interaction. The specific criteria formulated were that  $\hat{\omega}^2$  for one main effect was less than ten percent of the total  $\hat{\omega}^2$ , and further, was less than twenty-five percent of the  $\hat{\omega}^2$  for the other main effect. Such data was not predicted from the simple rules discussed earlier. Plotted in Figure 2(a), it corresponds to a diverging fan of straight lines, symmetrical about the horizontal axis. One interpretation of this category is that it represents multiplication of components in which the horizontal axis is coded unipolar and the vertical axis bipolar.

Category 6 - Significant Main Effects, Significant Interaction Not Concentrated in the Bilinear Term - "Curvilinear Multiplying Model"

Subjects in this category have two significant main effects, but tests of the interaction show a significant residual interaction after the bilinear term is extracted. This category, shown graphically in Figure 2(b), is clearly related to Category 3 (Figure 1(c)). The major distinction between them is a nonlinearity which is present in this model but not in the unipolar multiplying model.

Category 7 - "Unclassified"

Subjects in this category have data which do not meet the criteria of any of the categories discovered yet. Essentially they have relatively low total explained variance, yet they also have effects which reach significance.

Results

The results of the study are displayed in Table 1. For the adequacy-importance task only fourteen of the eighty-five subjects (16%) obeyed the assumptions of the adequacy-importance model, whereas forty-four of seventy-seven subjects (57%) obeyed the assumptions of the Fishbein model. This difference is significant ( $X^2 = 29.1$ , d.f. = 1, p < .001): Even under the conservative assumption of ordinal equivalence, which requires the adding model subjects of Category 2 to be combined with the unipolar multiplying models for the adequacy-importance task only<sup>7</sup>, the Fishbein model still performs significantly better ( $X^2 = 5.4$ , d.f. = 1, p < .05).

Subjects who performed the adequacy-importance task were more widely distributed over categories than those performing the Fishbein task, and particular interest settles upon the asymmetric multiplying and curvilinear multiplying model types. As discussed elsewhere (Bettman, Capon and Lutz, 1974) the curvilinear multiplying model results from subjects treating the importance component in a manner not indicated by the model. At low levels of belief, the affect value was initially low, rose, and then fell again as importance increased. There was no such effect at low levels of importance, indicating a possible ambiguity in the importance component. (Bettman, Capon, and Lutz, 1974, p. 17).

TABLE 1  
Results of Previous Study Comparing  
the Fishbein and Adequacy-Importance Models<sup>a</sup>

Category Number	Model Type	Adequacy-Importance Task	Fishbein Task
1	No Effects	3	4
2	Adding	19	18
3	Unipolar Multiplying	14	3
4	Bipolar Multiplying	11	44
5	Asymmetric Multiplying	13	4
6	Curvilinear Multiplying	18	2
7	Unclassified	<u>7</u>	<u>2</u>
TOTALS		85	77

<sup>a</sup>Adapted from Bettman, Capon, and Lutz, (1974).

The asymmetric model category is particularly interesting. One strong main effect and a strong interaction are required for inclusion, but the component which produces the main effect differs in type for the two tasks. For the adequacy-importance task, the adequacy component ( $B_{ij}$ ) demonstrates the strong main effect, while for the Fishbein task, the evaluation component ( $a_i$ ) exhibits the effect. (Within each task all subjects in the category, 13 for adequacy-importance, and 6 for Fishbein, had the same strong main effect).

This data pattern would be expected if, for each task, one of the two types of information was coded unipolar and the other bipolar, before multiplying. The unipolar dimensions from the two models were importance ( $I_i$ ) and likelihood ( $b_{ij}$ ), the bipolar dimensions being adequacy ( $B_{ij}$ ) and evaluation ( $a_i$ ). This bipolar coding for one component and unipolar coding for the other is contrary to the assumptions of both the adequacy-importance and Fishbein models.

Other "incorrect" coding occurred in the adequacy-importance task for eleven subjects who were classified as bipolar multiplying. Thus a total of 24 (11 + 13), or 28% of the subjects, treated the adequacy-importance data in a manner different from that assumed by the model on the coding dimension.

#### A New Model

The ambiguity of the importance component in the adequacy-importance model and the indication of possible ambiguity in the likelihood component of the Fishbein model suggest that a new model which is a hybrid of the two models may be superior to both of them.

This new model may be formulated as:

$$A_j = \sum_{i=1}^n B_{ij} a_i \quad (4)$$

where  $A_j$  = Attitude toward the brand;  $B_{ij}$  = Belief as to the extent to which attribute  $i$  is offered by brand  $j$ ;  $a_i$  = Evaluative aspect of attribute  $i$ , its goodness or badness; and  $n$  = the number of attributes. Both the evaluative and belief (adequacy) components are assumed to be coded in a bipolar manner and then multiplied. For purposes of further discussion, this new model will be referred to as the adequacy-evaluation model.<sup>8</sup>

The development of this model and the overall concern with validation of multi-attribute attitude models led to a second study, in which the aim was to test the adequacy-evaluation model by an ANOVA approach and to compare the results with those of the Fishbein and adequacy-importance tasks highlighted above. In addition, all three models were tested by both the individual correlation analysis approach and by cross-sectional correlation. It was hoped that the use of these three alternative validation procedures would result in greater insight into attitude structure than the single method approach employed in past studies.

#### A Multimethod Approach

##### Method

Sixty subjects from an undergraduate psychology class were each required to complete two tasks. In the first task, they were presented with a series of scales relating to different brands and attributes of toothpaste. They were asked to evaluate a series of attributes of toothpaste (on a good-bad scale) and then rate their beliefs as to the adequacy of certain brands in offering each attribute (on a very high-very low scale). Finally, they responded to an overall affect measure for each toothpaste brand (very favorable-very unfavorable). This data would be input into correlational tests of the adequacy-evaluation model.

The second task was identical to the ANOVA task described for the previous study. Subjects were presented with a series of profiles of the form:

You believe that Brand X is: very high :\_\_ :\_\_ :X :\_\_ :\_\_ : very low  
in possessing a quality which you personally believe is:  
very good :\_\_ :\_\_ :\_\_ :X : very bad.  
To you, using Brand X would be:  
very favorable :\_\_ :\_\_ :\_\_ :\_\_ :\_\_ :\_\_ :\_\_ :\_\_ :\_\_ :\_\_ : very unfavorable.

A total of fifty profiles of the hypothetical brand were rated, two replications each of the 5x5 factorial design. The instructions for the profile rating task emphasized that subjects were to rate a hypothetical product, Brand X, and not toothpaste; i.e., they were asked to consider how they would use information on attributes for products in general. Also, subjects were told that they should evaluate Brand X assuming that they had information only on the single attribute in the profile, and that the two pieces of information in the profile were intended to represent their own feelings about the attribute, not someone else's. The response scale used for the profiles was anchored (Anderson, 1974), by presenting subjects with "extreme" profiles first. For example:

Very high :X:\_\_:\_\_:\_\_: very low -- Very good :X:\_\_:\_\_:\_\_: very bad.

Order of the other profiles was randomized.

Data for input into the correlation analyses for the Fishbein and adequacy-importance models had been previously gathered in a similar manner from the subjects whose data were presented above in the discussion of the ANOVA results in the first study.

### Results

Three sets of results will be presented in this section. First, the adequacy-evaluation model is compared with the Fishbein and adequacy-importance models in the ANOVA framework; then the results from the individual correlation analyses and the cross-sectional correlation analyses are discussed.

ANOVA. Detailed results for the adequacy-evaluation model are shown in Table 2.<sup>9</sup> Comparison with similar data for the other two models (Bettman, Capon, and Lutz, 1974, p. 31) indicates a similar pattern of explained variance ( $\omega^2$ ) for each category. A comparison of how subjects are classified by the three models is presented in Table 3.

TABLE 2

Summary of ANOVA Results from the Adequacy-Evaluation Task

Category Number	Model Type	N	Adequacy ( $B_{ij}$ ) $\omega^2$	Evaluation ( $a_i$ ) $\omega^2$	Interaction $\omega^2$	$\Sigma\omega^2$	r
1	No Effects	1	.000	.000	.000	.000	-.367
2	Adding	20	.269	.460	.024	.753	.806
3	Unipolar Multiplying	4	.172	.326	.273	.771	.824
4	Bipolar Multiplying	23	.024	.054	.716	.794	.828
5	Asymmetric Multiplying	5	.033	.237	.439	.709	.759
6	Curvilinear Multiplying	5	.293	.254	.231	.778	.808
7	Unclassified	2	.337	.494	.049	.880	.889
TOTALS		60	.149	.253	.358	.760	.808

Once again a fairly large number of subjects, twenty, is classified into the adding category. However, twenty-three of the sixty subjects (38%) were classified into the "correct" bipolar multiplying category, a result which is inferior to the Fishbein task (57%) yet far superior to the adequacy-importance task (16%). Indeed, the pattern of data resembles that of the Fishbein task more than that of the adequacy-importance task, in that only a few subjects are classified into the unipolar, asymmetric and curvilinear multiplying categories. The conclusion from the ANOVA analysis is that of the three competing models, the Fishbein model is superior in capturing the judgment process of most subjects for the ANOVA rating task, but that the adequacy-evaluation model clearly outperforms the adequacy-importance model.

TABLE 3

## Comparison of ANOVA Results from Three Tasks

Category Number	Model Type	Adequacy-Evaluation Task	Adequacy-Importance Task	Fishbein Task
1	No Affects	1	3	4
2	Adding	20	19	18
3	Unipolar Multiplying	4	14	3
4	Bipolar Multiplying	23	11	44
5	Asymmetric Multiplying	5	13	4
6	Curvilinear Multiplying	5	18	2
7	Unclassified	<u>2</u>	<u>7</u>	<u>2</u>
TOTALS		60	85	77

Individual level correlation analysis. The results of the individual level correlation analyses are shown in Table 4. They are displayed by constructing a series of interval ranges for the correlation coefficient and placing subjects in the appropriate range. As shown in the table, 54% of the subjects in the adequacy-evaluation task demonstrated statistically significant correlations, while only 35% and 32% of the adequacy-importance and Fishbein task subjects did so. The average values of  $r$ , as computed through Fisher's  $r$  to  $z$  transformation, differed significantly between the adequacy-evaluation model and the Fishbein model ( $p < .01$ ), with the adequacy-importance model differing significantly from neither. Thus, based on the individual level correlation analysis, the Fishbein model appears to be inferior to the other two models.

Cross-sectional correlation analysis. The final set of analyses were cross-sectional correlations performed on the individual toothpaste brands. These results are presented in Table 5.

The first point of interest is the average amount of variance explained by each model, as computed from the adjusted  $r^2$  ( $r^2$  adj.) values, which indicates that the adequacy-evaluation model outperforms both the adequacy-importance model and the Fishbein model.

Further, for each individual brand, both the adequacy-evaluation and adequacy-importance models outperform the Fishbein model on every comparison.

TABLE 4

## Individual Level Correlation Results From Three Tasks

Range of Values of Correlation Coefficient (r)	Adequacy-Evaluation Task			Adequacy-Importance Task			Fishbein Task		
	n <sup>a</sup>	%	Cum %	n	%	Cum %	n	%	Cum %
.87-1.0 <sup>b</sup>	18	.32	.32	18	.22	.22	13	.18	.18
.75-.86 <sup>c</sup>	13	.23	.55	11	.13	.35	10	.14	.32
.41-.74	15	.26	.81	35	.42	.77	28	.37	.69
.21-.40	4	.07	.88	12	.15	.92	6	.08	.77
.00-.20	3	.05	.93	3	.04	.96	8	.11	.88
<.00	<u>4</u>	<u>.07</u>	1.00	<u>3</u>	<u>.04</u>	1.00	<u>9</u>	<u>.12</u>	1.00
TOTALS	57	1.00		82	1.00		74	1.00	
Average r		.759			.699			.596	

<sup>a</sup>Three subjects' data were eliminated from each task, due to constant values for either the predictor or criterion variable.

<sup>b</sup>Statistically significant correlation coefficient ( $p < .01$ ).

<sup>c</sup>Statistically significant correlation coefficient ( $p < .05$ ).

The adequacy-evaluation model exhibited higher  $r^2$  than the adequacy-importance model in four out of the six possible comparisons. The low correlations obtained for the Fishbein model in this case are somewhat perplexing; in all of the published work on the model, the relationships have seldom been as small as the ones observed here. Further, Nakanishi and Bettman (1974), in a study utilizing an instrument virtually identical to the one employed here, measured student attitudes toward the same seven brands of toothpaste and obtained an average  $r^2$  of .253, more than twice the size of the present findings. Therefore, the poor correlations found for the Fishbein model in this study can be somewhat discounted. This same conclusion can be applied to the results of the individual level correlation analysis as well, where the Fishbein model's performance was substantially weaker than in the Nakanishi and Bettman study. Obviously, the present findings may be more valid than the Nakanishi and Bettman results; however, the weight of other evidence for the Fishbein model dictates against this conclusion.

TABLE 5

## Cross-Sectional Correlation Results from Three Tasks

Toothpaste Brands	Adequacy-Evaluation Task (n=60)			Adequacy-Importance Task (n=85)			Fishbein Task (n=77)		
	ra	r <sup>2</sup>	r <sup>2</sup> adj.	ra	r <sup>2</sup>	r <sup>2</sup> adj.	r	r <sup>2</sup>	r <sup>2</sup> adj.
Ultradrite	.630	.397	.387	.442	.196	.186	.412 <sup>b</sup>	.170	.159
Pepsodent	.369	.136	.121	.559	.312	.304	.276 <sup>c</sup>	.076	.064
Macleans	.420	.176	.162	.477	.228	.219	.253 <sup>c</sup>	.064	.052
Crest	.632	.399	.389	.483	.233	.224	.239 <sup>c</sup>	.057	.045
Closeup	.665	.442	.433	.535	.286	.277	.474 <sup>b</sup>	.225	.215
Colgate	.619	.383	.373	.538	.289	.280	.245 <sup>c</sup>	.060	.048
Gleem	.505	.255	.242	.496	.246	.237	.377 <sup>b</sup>	.142	.131
Average r <sup>2</sup> adj.			.336			.281			.121

<sup>a</sup>All values statistically significant (p < .01).

<sup>b</sup>Significant (p < .01).

<sup>c</sup>Significant (p < .05).

## Discussion

This study reports a multimethod validation approach for three different multi-attribute attitude models. Two of the models tested were the Fishbein and adequacy-importance models, the third being a new model -- the adequacy-evaluation model -- whose genesis was suggested by the results obtained from an earlier study.

The three validation procedures employed were the familiar cross-sectional correlation, individual level correlation, and an ANOVA validation procedure new to the field of consumer behavior. The results which were obtained from these procedures were equivocal. From the ANOVA validation methodology, the adequacy-importance model was inferior to the other two models by virtue of the wide classification distribution of subjects. This did not occur to the same degree for the adequacy-evaluation and Fishbein models, but the Fishbein model outperformed the adequacy-evaluation model in terms of correct classifications by a considerable margin.

The individual correlation approach showed the Fishbein model to be inferior to the other two models, the adequacy-evaluation model being marginally superior to the adequacy-importance model by the criterion of  $r^2$ . The cross-sectional correlation analysis magnified the differences found at the individual level, the adequacy-evaluation model outperforming the adequacy-importance model, and both being superior to the Fishbein model.

Taking all of these results into consideration, it appears that the new adequacy-evaluation model is superior overall to both the adequacy-importance and the Fishbein models, for it does not perform badly on any one of the validation tests and is superior on two of the three.

Clearly this is not the last word in validation and multi-attribute models. Other models exist in the literature, and some of these should be subjected to a multimethod validation procedure. It is interesting to note that the genesis of the new model was due to a new validation procedure being applied at the individual level. Cross-sectional analysis modes might not have been able to isolate the possibility of such a model.

Finally, a note of caution is appropriate. The current findings are based on a single-attribute version of all three models tested. Thus the results may not be simply extended to encompass the multi-attribute case. Future research should extend the ANOVA approach to an examination of the additive assumption of multi-attribute models. In addition, other validation procedures should be employed and new models tested by a multimethod approach. Had a single approach been taken in this study, quite different conclusions might have been drawn, depending upon the procedure employed. Multimethod validation procedures allow the development of a richness in research findings which single methods cannot hope to achieve.

## FOOTNOTES

1. The authors gratefully acknowledge support of the Center for Marketing Studies at the Graduate School of Management, University of California, Los Angeles.
2. James R. Bettman, Noel Capon, and Richard J. Lutz are associate professor and assistant professors, respectively, at the Graduate School of



Management, University of California, Los Angeles. The ordering of the authors' names is alphabetical; all three contributed equally to the project.

3. To some extent the large number of judgments can be reduced by assuming that higher order interactions are negligible, in which case main effects and lower order interactions can be studied with fewer stimuli, either by eliminating within cell replications or by using fractional factorial designs (Anderson, 1972).
4. However, a great deal of work has been done on the beliefs-only model, where the power of  $a_i$  or  $I_i$  is assumed to be zero.
5. Some studies have taken disaggregative approaches and allowed each product term to be weighted by a regression coefficient.
6. It should be made clear that differences between the two models in terms of the coding schemes used to assign numerical values to scale responses are not inherent to the specific mathematical formulations shown in (1) and (2) above. Obviously, a researcher using data gathered under the Fishbein format could elect to code the scales from 1 to 5, rather than -2 to +2, and vice versa. In practice, however, research using the Fishbein model has treated the scales as bipolar, while adequacy-importance model research has coded the scales on a unipolar basis.
7. Green (1973) has shown that an interaction of the form shown in Figure 1(c) can be transformed through ANOVA to additivity, represented by Figure 1(a). Hence, the two patterns are "ordinally equivalent."
8. Strictly speaking, this "new" model is not really a new model so much as it is a new operationalization of the Fishbein model. Fishbein and Ajzen (1974) have suggested previously that a scale very similar to the adequacy measure used in this study may be a satisfactory operationalization of the  $b_{ij}$  construct.
9. Two measures of reliability of subjects' ratings are presented in Table 2;  $\Sigma \hat{\omega}^2$ , the total variance accounted for by the ANOVA, and  $r$ , the product-moment correlation between the two sets of twenty-five ratings.

#### REFERENCES

- Anderson, N.H. Application of an additive model to impression formation. Science, 1962, 138, 817-818.
- Anderson, N.H. Partial analysis of high-way factorial designs. Behavior Research Methods and Instrumentation, 1968, 1, 2-7.
- Anderson, N.H. Comment on 'An analysis of variance model for the assessment of configural cue utilization in clinical judgment.' Psychological Bulletin, 1969, 72, 63-65.
- Anderson, N.H. Functional measurement and psychophysical judgment. Psychological Review, 1970, 77, 153-70.

- Anderson, N.H. Algebraic models in perception. In E.C. Carterette and M. Friedman (Eds.), Handbook of perception. Vol. 2. New York: Academic Press, 1974, in press (a).
- Anderson, N.H. Cognitive algebra: Integration theory applied to social attribution. In L. Berkowitz (Ed.), Advances in experimental social psychology. New York: Academic Press, 1974, in press (b).
- Anderson, N.H. Information integration theory: A brief survey. In D.H. Krantz, R.C. Atkinson, R.D. Luce, and P. Suppes (Eds.), Contemporary developments in mathematical psychology. Vol. 2. San Francisco: W.H. Freeman, 1974, in press (c).
- Bass, F. and Wilkie, W.L. A comparative analysis of attitudinal predictions of brand preference. Journal of Marketing Research, 1973, 10, 235-241.
- Bettman, J.R., Capon, N. and Lutz, R.J. Cognitive algebra in multi-attribute attitude models. Working Paper No. 10, Center for Marketing Studies, University of California, Los Angeles, February 1974.
- Birnbaum, M.H. The devil rides again: Correlation as an index of fit. Psychological Bulletin, 1973, 74, 239-242.
- Bogartz, R.S. and Wackwitz, J.M. Polynomial response scaling and functional measurement. Journal of Mathematical Psychology, 1971, 8, 418-443.
- Brunswik, E. The conceptual framework of psychology. Chicago: University of Chicago Press, 1952.
- Brunswik, E. Perception and the representative design of experiments. Berkeley: University of California Press, 1956.
- Cohen, J.B., Fishbein, M. and Ahtola, O.T. The nature and uses of expectancy-value models in consumer attitude research. Journal of Marketing Research, 1972, 9, 456-60.
- Fishbein, M. An investigation of the relationships between beliefs about an object and the attitude toward that object. Human Relations, 1963, 16, 233-240.
- Fishbein, M., and Ajzen, I. Beliefs, attitudes, intentions, and behavior: An introduction to theory and research. Manuscript in press, 1974.
- Goldberg, L.R. Simple models or simple processes? Some research on clinical judgments. American Psychologist, 1968, 23, 483-496.
- Green, P.E. On the analysis of interactions in marketing data. Journal of Marketing Research, 1973, 10, 410-20.
- Hays, W.L. Statistics for psychologists. New York: Holt, Rinehart, and Winston, 1963.

- Hoffman, P.J. The paramorphic representation of clinical judgment. Psychological Bulletin, 1960, 57, 116-131.
- Hoffman, P.J., Slovic, P., Rorer, L.G. An analysis of variance model for the assessment of configural cue utilization in clinical judgment. Psychological Bulletin, 1968, 69, 338-349.
- Mazis, M.B., and Klippel, R.E. Instrumentality theories and consumer attitudes: Comparing alternative models. Proceedings. Association for Consumer Research, 1973, 346-7.
- Nakanishi, M., and Bettman, J.R. Attitude models revisited: An individual level analysis. Journal of Consumer Research, 1974, forthcoming.
- Rosenberg, M.J. Cognitive structure and attitudinal affect. Journal of Abnormal and Social Psychology, 1956, 53, 367-372.
- Sidowski, J.B. and Anderson, N.H. Judgments of city-occupation combinations. Psychonomic Science, 1967, 7, 279-280.
- Slovic, P. and Lichtenstein, S. Comparison of bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 7, 649-744.
- Wilkie, W.L. and Pessemier, E.A. Issues in marketing's use of multiattribute attitude models. Journal of Marketing Research, 1973, 10, 428-441.
- Wright, P. Research orientations for analysing consumer judgment processes. Proceedings. Association for Consumer Research, 1973, 268-79.

MULTI-ATTRIBUTE ATTITUDE MODELS: A COMPARATIVE ANALYSIS

Morris B. Holbrook and James M. Hulbert  
Columbia University

Existing research on attitude structure is summarized and a series of hypotheses derived. These hypotheses are tested on data gathered in the context of the 1972 Presidential election. Results suggest that overall attitude may be predicted quite well using beliefs and evaluations, with a limited role for salience in choosing appropriate attributes.

A consistent view of attitude has long eluded behavioral scientists. Indeed, recent research has tended to complicate, rather than simplify, the underlying issues. In this paper, we attempt to move closer to a resolution of these issues. First, we briefly review relevant literature summarizing different schools of thought in a series of models. From this analysis we derive a set of hypotheses which are then tested on a common, purposefully collected data base.

Research on the Underlying Structure of Attitude

Several psychologists (Rosenberg, 1956, 1960; Fishbein, 1963, 1965, 1967B, 1967C) have proposed a model of attitude toward an object ( $A_0$ ) as a function of two components ( $C_1$  and  $C_2$ ) where  $C_1$  is a set of cognitions (beliefs or perceptions) concerning the object and  $C_2$  the degree of affect (the evaluation or amount of satisfaction) associated with each cognition. In Fishbein's words, these models reflect:

The hypothesis that an individual's attitude toward any object is a function of his beliefs about the object and the evaluative aspects of those beliefs.  
(Fishbein, 1967B, p. 395)

Some attempt has often been made (Rosenberg, 1956; Fishbein, 1963, 1965) to select the  $n$  most salient cognitions, where 'salient' cognitions are assumed to be those which determine  $A_0$  (Fishbein, 1967B, p. 395; Cohen, Fishbein, & Ahtola, 1972, p. 457).  $A_0$  has then been postulated to be a function of the sum of the crossproducts of the  $n$   $C_1$  and  $C_2$  components. Or, algebraically:  $A_0 = f(\sum_{i=0}^n C_{1i} \cdot C_{2i})$ .

Sheth (May, 1972) has pointed out at least three crucial assumptions of such models: (1) two factors ( $C_1$  and  $C_2$ ) are needed to predict  $A_0$ ; (2) one  $C$  should be used to 'weight' the other in a multiplicative relationship; (3) the  $n$  weighted products should be summed without further weighting to form an overall prediction of  $A_0$  (for other statements of these assumptions, cf. Cohen, Fishbein & Ahtola, 1972, p. 117; and Talarzyk, 1972, p. 466).

Researchers working in the Rosenberg-Fishbein tradition have tested such models and found them to be satisfactory predictors of the criterion (Rosenberg, 1956; Bither & Miller, 1969; Hansen, 1969; Fishbein, 1963, 1965; Triandis & Fishbein, 1963; Fishbein & Hunter, 1964; Anderson & Fishbein, 1965; Sampson & Harris, 1970). Recently, however, several marketing applications of two-component models have concluded that an unweighted model of the form  $\Sigma C_1$  works about as well as (Sheth, May, 1972; Farley, et al., undated; Scott & Bennett, 1971; Cohen & Ahtola, 1971; Lehmann, 1971; Beckwith & Lehmann, 1973) or even better than (Sheth & Talarzyk, 1972; Moinpour & MacLachlan, 1971) the two-component  $\Sigma C_1.C_2$  version.

A troubling question thus arises: a well-established psychological finding has been repeatedly contradicted by studies in the marketing literature. Have psychologists erred in their formulation of the two-component model, or have marketing researchers failed to test the  $\Sigma C_1.C_2$  model in a theoretically appropriate manner?

This question was addressed by several articles in the Journal of Marketing Research (Cohen, Fishbein & Ahtola, 1972; Bass, 1972; Sheth, November, 1972; Talarzyk, 1972). In our view, however, this debate failed to deal with the problem in its fullest scope. Only a few of the relevant studies were discussed. In addition, the ways in which  $C_1$  and  $C_2$  have been measured in the studies cited above lead to the conclusion that there are three, not two, determinants of attitude being measured and that all three of these components are implicit in the Rosenberg-Fishbein formulations. These three components may be defined as follows:

(1) BELIEF ( $B_i$ ): the perceived extent to which some concept or object (o) is related to some other object i, some attribute i, some value i, or some goal i, (Fishbein, 1967A, p. 259; 1967B, p. 389)--as measured, for example, by a subjective probability or perceived instrumentality scale;

(2) EVALUATION ( $E_i$ ): the degree of favorability of affect--i.e., the attitude toward that other object i, attribute i, value i, or goal i (Fishbein, 1967A, p. 260)--as measured, for example, by a series of bipolar adjectival scales heavily loaded on the evaluative dimension of the Semantic Differential (Fishbein & Raven, 1967);

(3) SALIENCE ( $S_i$ ): the degree of importance of  $B_i$  and its associated  $E_i$  in determining the overall attitude ( $A_o$ ) toward the concept or object (o) (Fishbein, 1967B, p. 396)--as measured, for example, by a bipolar important-not important scale or by a ranking of the n object i's, attribute i's, value i's, or goal i's.

Obviously, when the three basic components B, E, and S, are reduced to a two-component  $\Sigma C_1.C_2$  model, different researchers may define those components differently. Indeed, examination of the studies cited above reveals that they have included at least four different ways of operationalizing  $C_1$  and  $C_2$ . It follows that these researchers were virtually certain to have reached different conclusions precisely because they were testing different models. We now examine these models individually.

The Rosenberg Model:  $A_o = \sum_{i=1}^n B_i \cdot F(E_i, S_i)$

Rosenberg(1956) proposed a two-component model in which  $C_2$  was called 'value importance' and was viewed as the importance of a value  $i$  as a source of satisfaction. Cohen, Fishbein, and Ahtola (1972), however, argued that "despite Rosenberg's use of the term 'value importance',  $V_i$  is not a measure of importance...but a measure of satisfaction or evaluation" (p. 456). In contrast, Sheth (November, 1972) quoted Rosenberg (p. 463) indicating that the latter was interested in the importance of values as defined above. Sheth concluded that "more or less importance of a valued state does not seem to mean the same thing as the evaluation of that valued state" (p. 463) and charged that "In their enthusiasm to relate other theories to Fishbein's, the authors have misunderstood and misinterpreted Rosenberg" (p. 462). There is some truth on both sides of this exchange, though neither seems to have recognized that Rosenberg's formulation of his second component relates to both the evaluation and the salience dimensions. Thus he confounds the components we labelled  $E_i$  and  $S_i$  within the same measure, which we designate  $F(E_i, S_i)$ . The resulting model of the form  $\sum B_i \cdot F(E, S)$  has been supported in a number of studies, some of which suggested that  $\sum B_i \cdot F(E, S)$  was a better predictor than perceived instrumentality alone (Rosenberg, 1956, 1960; Bither and Miller, 1969; Hansen, 1969).

The Fishbein Model:  $A_o = \sum_{i=1}^n B_i \cdot E_i$

Fishbein has tested a two-component model in which the  $C_1$  and  $C_2$  components correspond quite closely to what we have called  $B_i$  and  $E_i$ . A number of studies have reported good prediction with the  $\sum B_i E_i$  model (Fishbein, 1965; Sampson and Harris, 1970), while comparative research has supported the  $\sum B_i E_i$  model over competitors (Fishbein, 1963; Triandis and Fishbein, 1963; Fishbein and Hunter, 1964).

The Columbia Model:  $A_o = \sum_{i=1}^n B_i \cdot S_i$

A number of marketing studies have used two-component models of the form  $\sum B_i S_i$  to predict  $A_o$ . We call this model the 'Columbia Model' since its first application was apparently in a study of instant breakfast carried out at Columbia in the mid-1960's. It omits the  $E_i$  component, assuming implicitly that attributes are evaluated positively (with similar degree of affect).

Such applications have generally found that weighting  $B_i$  by the  $S_i$  component contributes little to prediction of  $A_o$  (Scott and Bennett, 1971; Hansen and Bolland, 1971; Sheth, May 1972). When  $B$  and  $BS$  scores have been used in regressions to predict the criterion, the  $S_i$  component has also been found to contribute little or nothing (Cohen & Ahtola, 1971; Sheth, May 1972; Farley, Howard and Weinstein, undated).

The Purdue Model:  $A_o = \sum_{i=1}^n F(B_i, E_i) \cdot S_i$

Several researchers--most of them associated in one way or another with Purdue--have employed measures of the first component which pertain to how satisfactory (Bass & Talaryk, 1972; Moinpour & MacLachlan, 1971; Kraft, et al., undated) or how ideal (Lehmann, 1971; Bass, et al., 1972; Beckwith & Lehmann, 1973) a brand is with respect

to a given attribute (i). Cohen, Fishbein, and Ahtola (1972) pointed out that such measures of  $C_1$  actually combine aspects of  $B_1$  and  $E_1$  into one component [which we designate  $F(B,E)$ ], concluding that the Purdue researchers have "proposed a new model which might be termed an 'adequacy-importance' model" (p. 456).

A number of studies have reported that  $\Sigma F(B,E)$  models predict about as well as  $\Sigma F(B,E).S$  (Bass and Talarzyk, 1972; Sheth and Talarzyk, 1972; Beckwith and Lehmann, 1973), thus reinforcing the Columbia findings that S-weights do not offer significant improvement in the prediction of overall affect.

### Some Recent Tests of Three-Component Models

The usefulness of S-weights is also challenged by recent psychological studies using all three components (B,E, and S) to predict  $A_0$ . Hackman and Anderson (1969) defined 'relevance' of a belief as "the degree to which a particular belief is important to a subject in evaluating an attitude object..." (p. 56). In predicting students' attitude toward a teacher who was rated on various attributes (i) (in turn rated on six-point scales for  $E_1$  and  $S_1$ ) no improvement of the  $\Sigma B.E.S$  over the  $\Sigma B.E$  model was found. A similar conclusion was reached in a study by Anderson (1970). He concluded that "the Fishbein formula, which considers the 'strength of belief' times the 'affect associated with that belief' does not benefit from the inclusion of another score, relevance of belief." (p. 46)

In summarizing studies using the  $\Sigma B.E.$  model which have, in one way or another, taken S-weights into account, Fishbein and Ajzen (1972) state the following rather free interpretation:

It has sometimes been argued that each piece of information should also be given a weight for its importance, salience, or relevance.... Despite the intuitive plausibility of this position, recent studies that have obtained measures of importance or relevance in addition to belief strength have consistently found that adding these weights to an expectancy-value model attenuates the prediction of attitude (p. 509; see also Cohen, Fishbein, and Ahtola, 1972, pp. 458-549).

### Conclusions on Predictive Power of 1-, 2-, and 3-Component Models

Not surprisingly we may conclude that the performance of various models in predicting  $A_0$  depends upon how their components are defined in the first place. In brief, when models of the form  $\Sigma B$  are compared with models of the form  $\Sigma B.E$  (Fishbein) or  $\Sigma B.F(E,S)$  (Rosenberg), the second component almost invariably contributes to the prediction of  $A_0$  (especially where the attributes are not all equally favorable). When, on the other hand, the comparison is between models of forms  $\Sigma B$  vs.  $\Sigma B.S$  (Columbia),  $\Sigma F(B,E)$  vs.  $\Sigma F(B,E).S$  (Purdue), or  $\Sigma B.E$  vs.  $\Sigma B.E.S$  (Three-Component Model), the addition of S-weights seems to contribute little predictive power.

This failure of S-weights in the Columbia, Purdue, and Three-Component Models may be explained in part by several important considera-

tions that have sometimes been overlooked in practice. First, we could question the Columbia Model's implicit assumption that  $E_i$  is more or less equally positive for all attributes used (cf. Bither & Shuart, p. 7). Second, most studies using the Columbia and Purdue models have consciously attempted to discover and use only the most salient attributes. If such attempts were successful, the inclusion of more or less equally important S-weights would clearly do little to improve prediction of  $A_0$ . Another problem in using S-weights arises if first-component scores are closely inter-related. If such multicollinearity exists separate S-weights can contribute little to prediction. Beckwith and Lehmann (1973) found such inter-correlations in their data, as did Farley, et al. (undated). An even more serious challenge to S-weights arises if the first and second components are somehow closely related. This type of redundancy would seem especially likely for the Purdue Model  $\Sigma F(B,E).S$  since there is strong a priori reason for thinking that  $E_i$  (the favorability of an attribute) might be strongly related to  $S_i$  (its importance). Finally, there has been a variety of objections to the methods used to measure the S-component (Sampson and Harris, 1970; Scott and Bennett, 1971; Alpert, 1971; Schendel, 1971). Perhaps most seriously, Sheth and Talarzyk (1972) argue that the whole procedure of rating attribute salience for a product class may be inappropriate if salience weights differ between brands, as indeed they might if the brands are not perceived as substitutes or if each brand is evaluated in terms of its 'strong points.'

#### Hypotheses About the Performance of 2- and 3-Component Models

We may summarize the considerations raised above by setting forth the following hypotheses as a way of reconciling the findings reported in the psychological and marketing literature.

First, we suggest that a model preferable conceptually to any of the two-component  $\Sigma C_1.C_2$  versions discussed above is the three-component model of the form  $\Sigma B.E.S$  which explicitly treats the B, E, and S components separately (cf. Hackmann & Anderson, 1970). Furthermore in testing such a model, attributes ( $i = 1, 2, \dots, n$ ) should be consciously chosen to display variation between attributes and across individuals in the  $B_i$ ,  $E_i$ , and  $S_i$  ratings. In other words, attributes should be selected not all of which are certainly possessed by the attitude object, not all of which are unequivocally favorable, and not all of which are of paramount salience for everyone in determining  $A_0$ .

Given the existence of these kinds of data, we might then check our interpretation of the empirical findings discussed above against the following predictions: (1) In accord with empirical findings supporting the Rosenberg Model,  $\Sigma B.(E.S)$  should perform better than  $\Sigma B$  in predicting  $A_0$ . (2) In accord with results supporting the Fishbein Model,  $\Sigma B.E.$  should perform better than  $\Sigma B$ . (3) In accord with repeated tests of the Columbia Model, the performance of the  $\Sigma B.S$  model should be quite low and should not be damaged by omitting the S-weights. (4) In accord with tests of the Purdue and Three-Component Models, when B or E and S are highly correlated and/or  $\text{Var}(B)$  or  $\text{Var}(E)$  and S are highly correlated,  $\Sigma B.E$  should perform about as well as  $\Sigma(B.E).S$  in predicting  $A_0$ . (5) With any given model, predictive power can always be improved by using multiple regression rather than



straight summation (neglecting the effects of a loss in degrees of freedom), but the above conclusions concerning the relative performance of B vs. BES, B vs. BE, B vs. BS, and BE vs. BES scores should still apply.

We note that, in reviewing the debate between Cohen, Fishbein and Ahtola; Bass; Sheth; and himself, Talarzyk (1972) suggested "the need for research to verify the similarities and differences between applications of models across areas of study" (p. 467). Although the study to be described below was designed and data collected before the appearance of Talarzyk's suggestion, we feel that it answers, in part, his call for research bearing on the relationship between the Rosenberg, Fishbein, Columbia, and Purdue Models.

### An Empirical Test

The above hypotheses--originally developed by Holbrook (1972)--were tested in the context of the 1972 presidential election. Just prior to the voting, 99 eligible student voters were asked to rate twenty-two attributes on (a)  $E_1$  ("how desirable you personally feel each attribute is in a presidential candidate" on a ten-point scale from "very undesirable" (scored -4.5) to "very desirable" (scored +4.5)) and (b)  $S_1$  ("how important you personally feel each attribute is in making your overall evaluation of a candidate" on a ten-point scale from "very unimportant" (1) to "very important" (10)). Immediately afterwards, the candidates--George McGovern and Richard Nixon--were rated on  $B_1$  for each attribute (in randomized order) according to "how likely you personally feel it is that \_\_\_\_\_ possesses each attribute" (using a ten-point scale from "very unlikely" (scored -4.5) to "very likely" (scored +4.5)). Then  $A_0$  and voting intention ( $I_0$ ) scores were obtained for each candidate on ten-point scales running from "dislike very much" (1) to "like very much" (10) and from "definitely will not" (1) to "definitely will" (10).

Informal inspection of the  $B_1$ ,  $E_1$ , and  $S_1$  scores suggested that we had succeeded in our attempt to include attributes representing a meaningful range of variation along these dimensions. For McGovern; mean B scores ranged from -3.17 (condones wiretapping) to 3.85 (favors a reduction in the national defense budget); their standard deviation varied from 1.23 (supports civil rights) to 3.00 (has a competent running mate). Similarly, for Nixon, B ranged from -3.89 (wants amnesty...) to 3.40 (believes in law and order), with SD of B varying from 1.31 (wants amnesty...) to 3.28 (...competent running mate). Mean E's ranged from -2.54 (condones wiretapping) to 4.20 (has good decision-making ability) with SD's of 0.96 (is honest) to 3.15 (wants amnesty...). Similarly mean S's ranged from 1.64 (attends church regularly) to 8.37 (...decision-making ability) with SD's of 1.33 (...decision-making ability) to 3.47 (condones wiretapping). (See Table 1.)

Correlation and regression analyses were used to test the hypotheses. Results are presented in Table 2.

Our first hypothesis is supported by the significant improvement of the  $\Sigma B.(E.S)$  over the  $\Sigma B$  model in predicting  $A_n$  (.735 vs. .490,  $Z = 2.88$ ,  $p = .002$ ),  $I_n$  (.691 vs. .428,  $Z = 2.70$ ,  $p = .003$ ),  $A_n$  (.785 vs.

TABLE 1  
Beliefs, Saliencies and Evaluations

Attribute	$\bar{B}_n$	Var ( $B_n$ )	$\bar{B}_n$	Var ( $B_n$ )	$\bar{S}$	Var (S)	$\bar{E}$	Var (E)
(1) Looks out for special interest groups	0.78	8.37	2.28	6.09	4.72	8.94	-1.10	8.06
(2) Supports the free enterprise system	1.11	7.29	2.76	5.24	6.62	4.69	2.24	5.28
(3) Supports civil rights	3.59	1.51	-0.82	6.24	7.36	5.80	3.12	4.32
(4) Has charisma	-0.63	6.78	-2.32	5.70	4.17	8.04	0.54	6.23
(5) Believes in law and order	1.73	6.35	3.40	3.33	6.27	6.69	2.10	5.54
(6) Condones wiretapping	-3.17	3.75	2.61	5.50	4.40	12.06	-2.54	5.37
(7) Favors a reduction in the national defense budget	3.85	2.46	-2.50	6.37	7.12	4.60	1.37	9.23
(8) Attends church regularly	1.00	6.56	-1.81	7.07	1.64	5.78	-.93	5.61
(9) Wants a national health insurance plan	3.00	2.29	-1.76	4.73	6.48	6.79	1.21	8.70
(10) Wants amnesty for draft resisters	3.26	3.49	-3.89	1.72	5.33	9.42	0.66	9.89
(11) Has good decision-making ability	0.50	8.77	1.64	7.71	8.37	1.78	4.20	1.04
(12) Promises to create jobs to reduce unemployment	3.46	1.87	-0.90	7.09	6.42	6.47	1.84	7.09
(13) Is gentle	1.81	4.83	-1.37	6.30	3.17	5.44	-.45	5.49
(14) Changes his mind frequently	1.81	6.03	0.04	7.80	5.17	6.91	-1.91	4.57
(15) Has a competent running mate	1.06	9.01	-1.39	10.78	7.23	3.75	3.04	2.33
(16) Wants universal price controls	0.38	7.80	-0.87	7.82	4.71	8.99	-1.03	7.16
(17) Favors normalized relations with Russia and Red China	2.37	5.70	2.82	5.33	7.42	3.71	3.07	3.32
(18) Is honest	2.22	5.90	-1.97	7.61	8.02	2.27	4.00	0.93
(19) Has an attractive wife	1.82	6.14	-0.79	9.51	2.11	6.62	-.77	5.78
(20) Is open and candid	1.56	8.04	-2.56	5.67	7.08	3.44	2.78	3.38
(21) Seeks consumer protection legislation	2.93	2.66	-1.00	5.30	6.91	4.83	2.95	3.18
(22) Wishes to close tax loopholes	3.58	1.93	-1.96	6.11	6.35	8.33	2.20	6.38

TABLE 2

Correlations of Various Models with Attitude ( $A_o$ ) and Intention ( $I_o$ )

	McGovern		Nixon	
	Attitude ( $A_m$ )	Intention ( $I_m$ )	Attitude ( $A_n$ )	Intention ( $I_n$ )
Simple r's with Summative Models:				
$\Sigma B$	.490	.428	.612	.613
$\Sigma B.S$	.648	.578	.699	.693
$\Sigma B.E$	.742	.700	.796	.784
$\Sigma B.E.S$	.735	.691	.785	.777
Multiple R's Using First Five Variables in Step-wise Regression				
B scores	.81	.80	.78	.78
BS scores	.82	.80	.78	.79
BE scores	.82	.77	.79	.79
BES scores	.81	.76	.79	.79

.612,  $Z = 2.51$ ,  $p = .006$ ), and  $I_n$  (.777 vs. 613,  $Z = 2.33$ ,  $p = .01$ ). Similarly, in support of our second hypothesis, the differences are of virtually the same magnitude and significance for comparisons between  $\Sigma B.E$  and  $\Sigma B$  models in predicting  $A_m$  (.742 vs. .490,  $Z = 2.88$ ,  $p = .002$ ),  $I_m$  (.700 vs. .428,  $Z = 2.82$ ,  $p = .002$ ),  $A_n$  (.796 vs. .612,  $Z = 2.71$ ,  $p = .003$ ), and  $I_n$  (.784 vs. .613,  $Z = 2.34$ ,  $p = .01$ ). Clearly, then, both the Rosenberg-like and the Fishbein two-component versions are--as usual--better predictors than a naive one-component form.

Turning to those models which use S-weights in connection with a first component represented by B or (B.E), we find that--contrary to our third hypothesis--the  $\Sigma B.S$  model does offer some partially significant tendency toward improvement over  $\Sigma B$  in predicting  $A_m$  (.648 vs. .490,  $Z = 1.62$ ,  $p = .053$ ),  $I_m$  (.578 vs. .428,  $Z = 1.40$ ,  $p = .081$ ),  $A_n$  (.699 vs. .612,  $Z = 1.10$ ,  $p = .136$ ), and  $I_n$  (.693 vs. .613,  $Z = .97$ ,  $p = .166$ ). This mixed disconfirmation of our hypothesis may be explained by the high correlations between E and S (a) for individuals across attributes ( $r_{ES} = .618$ ,  $p < .001$ ), (b) for attributes across individuals ( $r_{ES} = .519$ ,  $p < .001$ ), and (c) for  $E_1$  and  $S_1$  across attributes ( $r_{E_1 S_1} = .83$ ,  $p < .001$ ). These strong relations between E and S may have meant that S was in effect acting as a proxy for E in the  $\Sigma B.S$  model thereby imparting to it some of the superiority of the  $\Sigma B.E$  over the  $\Sigma B$  model. At any rate, this high degree of redundancy between E and S establishes the conditions under which we raised the fourth hypothesis. In accordance with this

hypothesis, we find that  $\Sigma B.E$  performs about as well as  $\Sigma(B.E).S$  in predicting  $A_m$  (.742 vs. .735, not significant),  $I_m$  (.700 vs. .691, n.s.),  $A_n$  (.796 vs. .785, n.s.), and  $I_n$  (.784 vs. .777, n.s.). These results support the contention that, when E and S are highly correlated, the  $\Sigma B.E.S$  does not improve upon the  $\Sigma B.E$  model.

Our final hypothesis was supported in part by the improvement over the  $\Sigma B$  model of the five-variable stepwise regression using B scores to predict  $A_m$  (.81 vs. .49,  $Z \approx 4.04$ ,  $p \lesssim .001$ ),  $I_m$  (.80 vs. .428,  $Z \approx 4.36$ ,  $p < .001$ ),  $A_n$  (.78 vs. .612,  $Z = 2.30$ ,  $p = .011$ ) and  $I_n$  (.78 vs. .613,  $Z = 2.30$ ,  $p = .011$ ). Similarly, stepwise regression with the BS scores tended to improve over the  $\Sigma B.S$  model in predicting  $A_m$  (.82 vs. .648,  $Z \approx 2.61$ ,  $p \approx .005$ ),  $I_m$  (.80 vs. .578,  $Z \approx 2.99$ ,  $p = .001$ ),  $A_n$  (.78 vs. .699,  $Z \approx 1.22$ ,  $p = .111$ ), and  $I_n$  (.79 vs. .693,  $Z = 1.53$ ,  $p = .063$ ). (In these comparisons Z-values are only approximate though some attempt has been made to allow for the loss of degrees of freedom in the stepwise regression.)

In the case of the BE and BES scores, the summative models performed well enough so that the improvements gained from multiple regression were marginal at best (and in no case significant, even at  $p = .10$ ). Nor was there any support for our anticipation of a difference in favor of regression models using BE or BES scores. Regressions with B scores performed fully as well as any competing models.

In summary, we might conclude that, for our data, the  $\Sigma B$  model performs less well than the  $\Sigma B.E$  or  $\Sigma B.E.S$  models, with the  $\Sigma B.S$  model falling somewhere in between, perhaps because S serves as a proxy for E. In addition, maximal predictions can be obtained in stepwise regression using as few as five B scores.

#### Is There a Role for Salience?

What then, if anything, is left for the role of salience (S)? In our data, the sizeable redundancy between E and S obviated any major contribution of the latter variable. It should be possible, however, to remove the redundancy between E and S by conjuring up attributes which are highly important, yet highly unfavorable (dishonesty, inability to make good decisions, opposition to civil rights). Presumably such salient attributes--though now expressed negatively--would still be important as determinants of  $A_o$ .

Indeed, it appears that the most useful role of salience might lie in selection of a limited set of attributes to use in predicting  $A_o$ . (The findings of Fishbein and his colleagues discussed earlier are not inconsistent with this conclusion. Some of their best predictions were obtained with attributes selected for their general or idiosyncratic salience). We noticed, for example, that the first one or two attributes to enter the stepwise regressions (decision-making and civil rights for McGovern, candidness and civil rights for Nixon) were also among the highest in saliency. In fact, when the individual correlations between B scores and  $A_o$  for each attribute were correlated with the attribute's S-weights, r's were found of .58 ( $p = .002$ ) for McGovern and .44 ( $p = .022$ ) for Nixon. This relation between the salience of a belief and its ability to predict  $A_o$ , however, might

(in light of the close relation between E and S) be interpreted as nothing more than a reflection of the almost tautological relation between a belief scale's E-value and its relation to  $A_0$  (i.e.,  $r_{r_{B_i} A_0} = .81$  for McGovern and  $.78$  for Nixon) due to the tendency of these correlations to shift from negative to positive as  $E_i$  varies between negative and positive valences. In an attempt to clear up some of this ambiguity introduced by the directionality of  $E_i$ , we therefore looked at  $(B.E)_i$  scores for which higher values should unambiguously predict  $A_0$ . The correlation between BE and  $A_0$  increased dramatically with an attribute's salience (S) ( $r_{r_{BE_i} A_0} = .74$  for McGovern and  $.56$  for Nixon--both significant at  $p = .003$ ). By contrast, the relation of these correlations to E itself fell to  $r = .47$  for McGovern and  $.56$  for Nixon. In addition, when  $r_{BE_i A_0}$  was regressed on S and E together, only S attained significant regression coefficients ( $t = 4.27$ ,  $p < .001$  for McGovern;  $t = 2.53$ ,  $p < .02$  for Nixon). In our view, this latter finding warrants the tentative conclusion that, when the direction of E is taken into account (eg., by weighting  $B_i$  by  $E_i$  or simply using only attributes with positively evaluated E's), the relation between B and  $A_0$  increases in strength with attribute salience (S). More specifically, it appears that good predictions of  $A_0$  using B scores could be obtained using only those few attributes for each individual which are (a) positively evaluated and (b) highly salient.

This conclusion could not be tested directly since many subjects assigned a great number of attributes the maximum S and E ratings. As a partial test, however, we selected the five attributes (numbers 3, 11, 15, 17 and 18) which were on the average judged highest in both S and E and were among the seven lowest attributes in Var (E). Using only these five attributes in a summative model ( $\Sigma B$ ), predictions of  $A_0$  were obtained ( $r = .70$  for McGovern and  $.71$  for Nixon--with 94 usable questionnaires) which rivaled those of the full  $\Sigma B.E.S$  model with all 22 attributes ( $r = .735$ ,  $Z = .08$ , n.s., and  $.785$ ,  $Z = 1.28$ , n.s., respectively). Moreover, when only the  $n$  attributes of these five which were (a) above the average in E and (b) at the highest level of S for any given subject were used in a summative model ( $\Sigma B/n$ ) to predict  $A_0$ , slightly better predictions were obtained for both McGovern ( $r = .730$ ) and Nixon ( $r = .769$ ). The average number of attributes ( $n$ ) used in these predictions was only 2.7 (median = 3). These results suggest, then, that--using as few as the three most salient (positively evaluated) attributes for each individual in a  $\Sigma B$  model--predictions of  $A_0$  may be achieved that are as good as those obtained from the full array of twenty-two attributes in the  $\Sigma B.E.S$  version, which in turn performs about as well as a model using B, BE, or BES scores in stepwise regression. (The differences between the  $\Sigma B/n$  and five B score regression models are not significant at  $p = .10$ ). We may conclude, then, that in the data described above individual salience weights are useful in that they may be used to specify for each respondent a very limited number of, say, three positively evaluated attributes whose B scores may be used in a simple summative model which predicts  $A_0$  virtually as well as models using far more information yet attaining only marginal improvements in performance.

These findings corroborate the results of Wilkie and Weinreich (1973). They compared the performance of a variety of summative models and found that when the order of inclusion of weights was the

same for all subjects, maximum correlation was attained with five attributes. When the order of inclusion was determined idiosyncratically, best fits occurred with only three attributes.

### Concluding Comments

Attribute models of attitude structure have received a great deal of marketing research effort in recent years. By now, a number of useful generalizations appear finally to have emerged. In this connection, the results of this study are strikingly similar to those reported by Wilkie and Weinreich (1973); moreover they relieve some of the concerns expressed by those authors.

First, since the  $\Sigma B$  predictions based on an aggregate selection procedure (high  $\bar{S}$  and  $\bar{E}$ , low Var (E)) were not significantly lower than the  $\Sigma B/n$  predictions based on idiosyncratic attribute selection, it appears that, for applied marketing research, prior selection of a relatively small set of attributes based on highest  $\bar{S}$ 's (or some similar criterion) can be relied upon to result in good prediction.

Second, where the analyst is interested in the consumer's belief structure (eg. as a variable intervening between message exposure and affect), the results suggest that selection of the three idiosyncratically highest- $\bar{S}$  attributes provides a satisfactory measure.

Third, the findings of this study and those of Wilkie and Weinreich suggest that the Purdue and Columbia practices of using a limited number of salient attributes is justified using the pragmatic criterion of prediction of  $A_c$  (though in the Columbia case this conclusion only applies if attributes with positive affect are used).

Finally, the present study used a large set of twenty two attributes, varying widely in B, E and S. The similarity of our results to theirs should alleviate Wilkie and Weinreich's expressed concern over their use of only seven attributes.

### FOOTNOTES

1. The authors wish to express their gratitude to Georgi Zahariev and Bernardo Cohen for their help in the early stages of the project and to Harry Steinberg, for his invaluable assistance in later stages. They also wish to thank the Faculty Research Fund of Columbia University Graduate School of Business for its support of the project.
2. Morris B. Holbrook is a doctoral candidate at the Columbia University Graduate School of Business. James M. Hulbert is Associate Professor at the Columbia University Graduate School of Business.

### REFERENCES

- Alpert, M.I. Identification of determinant attributes: A comparison of methods. Journal of Marketing Research, 1971, 8, 184-191.

- Anderson, L. R. Prediction of negative attitude from congruity, summation and logarithmic formulae for the evaluation of complex stimuli. Journal of Social Psychology, 1970, 81, 37-48.
- Anderson, L. R. & M. Fishbein. Prediction of attitude from the number, strength and evaluative aspect of beliefs about the attitude object: A comparison of summation and congruity theories. Journal of Personality and Social Psychology, 1965, 3, 437-433.
- Bass, F. M. Fishbein and brand preference: A reply. Journal of Marketing Research, 1972, 9, 461.
- Bass, F. M., E. A. Pessemier & D. R. Lehmann. An experimental study of relationships between attitudes, brand preference and choice. Behavioral Science, 1972, 17, 532-541.
- Bass, F. M. & W. W. Talarzyk. An attitude model for the study of brand preference. Journal of Marketing Research, 1972, 9, 42-96.
- Beckwith, N. E. & D. R. Lehmann. The importance of differential weights in multiple attribute models of consumer attitude. Journal of Marketing Research, 1973, 10, 141-145.
- Bither, S. W. & S. J. Miller. A cognitive theory view of brand preference. Proceedings, American Marketing Association, Fall, 1969, 280-286.
- Bither, S. W. & A. H. Shuart. Uses and misuses of expectancy attitude models in marketing. Pennsylvania State University, unpublished paper, undated.
- Cohen, J. B. & O. T. Ahtola. An expectancy X value analysis of the relationship between consumer attitude and behavior. Proceedings, Second Annual Conference Association for Consumer Research, 1971, 344-364.
- Farley, J. U., J. A. Howard & D. Weinstein. The relationship of liking and choice of an alternative to its attributes and their saliency. Columbia University, unpublished paper, undated.
- Fishbein, M. An investigation of the relationships between beliefs about an object and the attitude toward that object. Human Relations, 1963, 16, 233-249.
- Fishbein, M. Prediction of interpersonal preferences and group member satisfaction from estimated attitudes. Journal of Personality and Social Psychology, 1965, 1, 663-667.
- Fishbein, M. A consideration of beliefs and their role in attitude measurement. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: John Wiley and Sons, 1967a.
- Fishbein, M. A behavior theory approach to the relations between beliefs about an object and the attitude toward the object. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: John Wiley and Sons, 1967b.

- Fishbein, M. Attitude and the prediction of behavior. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: John Wiley and Sons, 1967c.
- Fishbein, M. & I. Ajzen. Attitudes and opinions. In P. H. Mussen and M. R. Rosenzweig, (Eds.), Annual review of psychology. Palo Alto: Annual Review, 1972. Pp. 487-544.
- Fishbein, M. & R. Hunter. Summation versus balance in attitude organization and change. Journal of Abnormal and Social Psychology, 1964, 69, 505-510.
- Fishbein, M. & B. H. Raven. The AB scales: An operational definition of belief and attitude. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: John Wiley and Sons, 1967.
- Hackman, J. R. & L. R. Anderson. The strength, relevance and source of beliefs about an object in Fishbein's attitude theory. Journal of Social Psychology, 1968, 76, 55-67.
- Hansen, F. Consumer choice behavior: An experimental approach. Journal of Marketing Research, 1969, 6, 436-443.
- Hansen, F. & T. Bolland. The relationship between cognitive models of choice and non-metric multidimensional scaling. Proceedings, Second Annual Conference, Association for Consumer Research, 1971, 376-388.
- Holbrook, M. B. The definition, measurement and prediction of attitude. Columbia University, unpublished paper, 1972.
- Kraft, F. B., D. H. Granbois & John O. Summer. Brand evaluation and brand choice: A longitudinal study. Indiana University, unpublished paper, undated.
- Lehmann, D. R. Television show preference: Application of a choice model. Journal of Marketing Research, 1971, 8, 47-53.
- Moinpour, R. & D. L. MacLachlan. The relationship among attribute and importance components of Rosenberg-Fishbein type attitude models: an empirical investigation. Proceedings, Second Annual Conference, Association for Consumer Research, 1971, 365-375.
- Rosenberg, M. J. Cognitive structure and attitudinal affect. Journal of Abnormal and Social Psychology, 1956, 53, 367-372.
- Rosenberg, M. J. An analysis of affective-cognitive consistency. In C. I. Hovland and M. J. Rosenberg (Eds.), Attitude organization and change. New Haven: Yale university Press, 1960.
- Sampson, P. & P. Harris. A user's guide to Fishbein. Journal of the Market Research Society, 1970, 12, 145-166.
- Schendel, D. E., W. L. Wilkie, and J. M. McCann. An experimental investigation of attitude importance. Proceedings, Second Annual Conference, Association for Consumer Research, 1971, 404-416.



- Scott, J. E. & P. D. Bennett. Cognitive models of attitude structure: Value importance is important. Proceedings, American Marketing Association, 1971, 346-350.
- Sheth, J. N. Measurement of attitude scores from beliefs and importances. College of Commerce and Business Administration. University of Illinois. Faculty working paper no. 50, 1972.
- Sheth, J. N. Reply to comments on the nature and uses of expectancy-value models in consumer attitude research. Journal of Marketing Research, 1972, 9, 465-467.
- Sheth, J. N. & W. W. Talarzyk. Perceived instrumentality and value importance as determinants of attitudes. Journal of Marketing Research, 1972, 9, 6-9.
- Talarzyk, W. W. A reply to the response to Bass, Talarzyk and Sheth. Journal of Marketing Research, 1972, 9, 465-467.
- Triandis, H. C. & M. Fishbein. Cognitive interaction in person perception. Journal of Abnormal and Social Psychology, 1963, 67, 446-453.
- Wilkie, W. L. & R. P. Weinreich. Effects of the number and type of attributes included in an attitude model: More is not better. Krannert Graduate School of Industrial Administration, Purdue University. Working paper no. 385, 1973.

PREDICTION OF ATTITUDES:  
A COMPARATIVE STUDY OF THE  
ROSENBERG, FISHBEIN AND SHETH MODELS

Secil Tuncalp and Jagdish N. Sheth<sup>1</sup>  
University of Illinois

Predictive efficacies of the Rosenberg, Fishbein and Sheth models of attitudes were tested in a comparative study of predicting consumer's attitudes toward brands of shampoo. Four different criteria of predictive efficacy were utilized. (i) Generalizability--variability of predictions across different attitudinal objects; (ii) Consistency--variability of predictions across different measures of attitudes; (iii) Stability--variability of predictions across different samples; (iv) Reliability--variability of predictive validations. Without a single exception, the Sheth model produced significantly better correlations with consumer's attitudes toward hair shampoos on each of the above four criteria. The better predictions with the Sheth model are attributed to truncation of scales, retaining multidimensionality of beliefs and more realistic wording of the questions.

Most researchers agree that attitudes refer to that mental state of the individual which represents his positive, negative or neutral feelings toward an object, concept or idea. Attitudes have become an important area of study in psychology largely because of a widely held belief that they precede the individual's behavior toward the object or concept, and hence can be used as important predictors of behavior. Whether or not attitudes precede behavior is a moot question. However, considerable attention has been given to theorize about the determinants of an individual's attitude (positive or negative feelings) toward an object or concept and how attitudes change over a period of time.

In the early history of attitudes, most of the research can be categorized as essentially definitional and descriptive. A new phase emerged in attitude research with the publication of Thurstone's work (1928) which attempted to quantitatively measure attitudes and to psychometrically scale attitudes toward specific objects such as the church. Considerable imaginative research has followed since Thurstone in which scholars in social psychology have attempted to build operational models of attitude structure by conceptualizing about factors which determine a person's attitude toward an object or concept (See Bem 1970; Fishbein 1967; McGuire 1969; and Triandis 1971).

Among the more popular models are the Rosenberg (1956, 1960) and the Fishbein (1963, 1967) models of attitudes. According to Cohen, Fishbein and Ahtola (1972), these models are part of a generic class of attitude models based on cognitive psychology which presume an expectancy-value formulation of attitudes: an individual has a positive or negative feeling (attitude) toward an object or concept because of his cognitive expectations about the object's capability to do certain things and the perceived values he attaches to those things. Unfortunately, there is considerable controversy in social psychology about the form and superiority of various models of attitudes (Rosenberg 1968). In consumer behavior, research on attitudes can be classified into three categories: First, the bulk of research consists of direct and sometimes indiscriminate applications of the attitude models developed in social psychology. This is especially true of the Rosenberg and the Fishbein models which have been applied to predict brand preferences and brand choice behavior (Wilkie and Pessemier 1973). To us it appears that this type of research suffers from the same problems we experienced in the fifties and early sixties when personality theories and stochastic models were directly applied in consumer behavior. Second, some researchers have consciously attempted to reformulate and adapt the expectancy-value models to make them suitable to consumer behavior (Bass and Wilkie 1973; Cohen and Ahtola 1971; Day 1971; Raju and Sheth 1974; Sheth and Talarzyk 1972; Sheth 1973). Finally, only a handful of researchers in consumer behavior have formulated their own models of attitudes by working backwards from consumer behavior to behavioral sciences and judiciously choosing and integrating the diverse concepts and theories of attitudes in pure disciplines (Banks 1950; Howard and Sheth 1969; Sheth 1969, 1971, and 1974; Myers and Alpert 1968; Alpert 1971). The dominant characteristic of these indigenous models is lack of allegiance to any specific school of thought or to a specific model in social psychology.

In an applied discipline, the researcher is often on the horns of dilemma: should he extend a model developed and tested in other discipline to his area or should he develop his own model? It is relatively easy to apply other people's models especially if those models are operationalized and measurement instruments are available as is very true with personality tests. On the other hand, there is always that nagging feeling whether the models from other disciplines are relevant at all, and if so, how much. In consumer behavior, we have witnessed similar dilemmas in the areas of personality research and mathematical modeling of brand loyalty (Sheth 1974a). Often the only way to get off the dilemma is to carry out comparative research and measure the efficacies of the borrowed and the indigenous models. Our objective in this study, accordingly, is to investigate the relative efficacies of the borrowed and the indigenous models of attitudes. Specifically, this study examines the predictive efficacy of the Rosenberg and the Fishbein models of attitudes which represent the popular borrowed models, and the Sheth model of attitudes which represents the indigenous model in consumer behavior.

Predictive efficacy of attitude models can be measured in at least four different ways. First, does one model in a comparable situation perform consistently better over other models across

different attitudinal objects? This is a test of the generalizability of the superiority of one model over other models. Second, does the model perform consistently better over other models across different but highly related measures of attitudes? In other words, if we measured the individual's positive or negative feeling (attitude) toward an object in several different ways, does one model consistently predict better over other models? Third, does one model perform better over others across different samples of observations? Specifically, how stable are the models when subjected to random sampling errors by way of split half sampling procedures? Finally, does one model consistently predict better than others when applied to a new sample? In other words, what is the reliability of model prediction across different samples?

The objective of this study is to examine the relative strengths and weaknesses of the Fishbein, the Rosenberg and the Sheth models in predicting consumer attitudes toward various brands of shampoo by utilizing the above four criteria of predictive efficacy, namely the generalizability, consistency, stability and reliability of predictions.

#### Brief Description of Three Attitude Models

The Rosenberg's model (1956, 1960) of attitudes theorizes that an individual's positive or negative feelings toward an object or concept is a function of two cognitive factors: (a) the perceived instrumentality (PI) of that object to block or attain a set of valued states, goals or objectives, and (b) the relative importances (VI) of those valued states, goals or objectives to the individual. Operationally, the model is stated as follows:

$$A_o = \sum PI_i X VI_i$$

Where  $A_o$  = Attitude toward the object or concept

$PI_i$  = Perceived instrumentality of the object or concept in blocking or attaining the  $i$ th valued state, goal or objective

$VI_i$  = Importance of  $i$ th valued state, goal or objective

While Rosenberg has limited his own research to only those valued states which reflect fundamental human values and objectives (individual freedom, procreation, economic well-being, etc.), most researchers in consumer behavior have applied the model to predict brand preference by defining product benefits as valued states and measuring the perceived instrumentality of a brand in providing or blocking those product benefits (Hansen 1972; Sheth and Talarzyk 1972; Sheth 1973).

The Fishbein model (1963, 1967) of attitudes toward object ( $A_o$ ) also presumes that an individual's positive or negative feeling toward an object or concept is a function of two factors: (a) the personal beliefs of the individual about the object or concept possessing or

or not possessing certain characteristics or attributes ( $B_i$ ) and (b) the personal evaluations of those characteristics or attributes ( $a_i$ ). Operationally, the Fishbein model is stated as follows:

$$A_o = \sum B_i a_i$$

Where  $A_o$  = Attitude toward the object or concept

$B_i$  = Personal belief about the probability of the object possessing the  $i$ th attribute

$A_i$  = Personal evaluation of the  $i$ th attribute

The two models are strikingly similar in their operationalization. While there is some controversy as to whether the Rosenberg and the Fishbein models are conceptually equivalent (Sheth 1972; Sheth and Park 1973), the Fishbein model has been more widely borrowed and applied in consumer behavior (Wilkie and Pessemier, 1973).

The Sheth model (1969, 1971, 1974) theorizes that attitude toward a brand is a function of perceived brand potential to satisfy consumer's needs, wants and desires related to the product consumption. The perceived brand potential is assessed by the consumer's evaluation of the brand on a set of criteria which link specific product attributes and the consumer's buying needs or motives. The profile of perceived brand potential is then reduced to a vector of evaluative beliefs ( $EB_i$ ) utilizing the rank reduction techniques in psychometrics. The vector of beliefs represents the consumer's assessment of the brand potential as a point in a multidimensional space of buying criteria. The number of dimensions is presumed to vary with product categories. Operationally, the Sheth model of attitudes is stated as follows:

$$A_o = \beta_1 EB_1 + \beta_2 EB_2 + \dots + \beta_n EB_n + e$$

Where  $A_o$  = Attitude toward the brand

$EB_{1,2,\dots,n}$  = Dimensions of evaluative beliefs

$\beta_{1,2,\dots,n}$  = Parameters to be estimated, and  $e$  = error

Conceptually and operationally, there are several differences between the Sheth and the Rosenberg-Fishbein-type expectancy-value models of attitudes. First, the cognitive beliefs underlying attitudes in the Sheth model are restricted to only those aspects of the brand which represent potential of the brand to satisfy consumer's buying and consuming motives. In other words, the object or concept is looked upon strictly as a goal-object based on the functional and the learning theories of attitudes. The Rosenberg model, at least in social psychology, has always dealt with more fundamental and non-product specific goals or valued states. While Fishbein's model is specific to the object with respect to the beliefs, the beliefs are not limited to only those beliefs which represent linkage between product attributes and consumer's motives. In this study, however, we have utilized comparable sets of beliefs among the three models in

view of the fact that researchers in consumer behavior have liberalized the Rosenberg and the Fishbein models to focus on the brand as a goal-object where the goals are buying and consuming motives.

Second, both the Rosenberg and the Fishbein models explicitly weight each belief respectively by its relative importance and evaluation. Both the models have, therefore, two cognitive components and operationally the relationship between the beliefs and the values are defined as inter-active or multiplicative. In the Sheth model, the cognitive beliefs are limited to the brand potential profile or the assessment of the brand's capability to satisfy a number of buying criteria. Of course, the one component Sheth model will be less costly in terms of data collection because it is more parsimonious.

Third, and perhaps the most significant difference between the expectancy-value models borrowed from social psychology and the indigenous model is with respect to the assumptions about the dimensionality of the cognitive structure underlying a person's attitude toward an object or concept. Both Rosenberg and Fishbein sum the weighted beliefs and thereby presume that the cognitive structure underlying attitudes is unidimensional. This is parallel to Thurstone's thinking about attitudes. On the other hand, the Sheth model presumes multidimensionality of cognitive structure allowing for the number of underlying dimensions to vary from product class to product class. If the consumer is satisfying a single motive, the Sheth model presumes that the dimensionality of the cognitive beliefs will be reduced to one. The number of dimensions is, therefore, clearly a function of the complexity and multiplicity of buying motives. More the number of uncorrelated and independent motives in a buying situation, the greater will be the dimensionality of the cognitive structure underlying attitudes.

A fourth difference between the Rosenberg-Fishbein type of expectancy-value models and the Sheth model relates to the question of bipolar scales. The Sheth model explicitly rejects the notion of true bipolarity of cognitive beliefs on the ground that such true bipolarity does not exist in most product benefits and that it unnecessarily creates skewness of distribution in the data making any statistical analysis more difficult. The scales in the Sheth model are, therefore, truncated to a range which seems realistic.

The fifth and the final difference relates to the usefulness of these models to public policy or industry decision makers. In the expectancy-value models of the Rosenberg and Fishbein type, it is not possible for the decision maker to identify a particular belief which he should control, manipulate or change in order to bring about attitude change. This is because the beliefs are aggregated into a single sum score. While this sum score may be very strongly related to a person's attitude toward an object, it is impossible to identify the relative contribution of a specific belief in determining a person's attitude. In the Sheth model, this problem does not exist. After reducing the brand potential profile to its underlying dimensions, the Sheth model directly relates each type of beliefs to the overall







## HEAD AND SHOULDERS

leaves scalp nondry            improbable \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:probable

Fishbein's evaluative component:

Lots of lathering            good \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:bad  
 Manageable hair            good \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:bad  
 Dandruff control            good \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:bad  
 Clean hair                  good \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:bad  
 Conditioned hair            good \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:bad  
 Soft hair                    good \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:bad  
 Pleasant smell              good \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:bad  
 Nondry scalp                good \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:bad

Sheth's evaluative belief component:

Q. Listed below are several scales constructed from product attributes most people use to judge the quality of hair shampoos. Please evaluate HEAD AND SHOULDERS (relative to all other brands) on each of the following criteria assuming that you are about to purchase it.

Makes lots of lather \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:Makes not much lather  
 Leaves hair manageable \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:Leaves hair unmanageable  
 Poor dandruff control \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:Excellent dandruff control  
 Leaves hair squeaky clean \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:Leaves residue and  
 does not rinse completely  
 Poor hair conditioner \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:Good hair conditioner  
 Leaves hair coarse \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:Leaves hair soft  
 Smells pleasant \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:Smells unpleasant  
 Leaves scalp dry \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:Leaves scalp oily

In addition, three different measures of attitude toward the object were obtained to examine the question of consistency.

Q. Please indicate the extent to which you think HEAD AND SHOULDERS is good or bad.

In general, HEAD AND SHOULDERS is very good \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: In general, HEAD AND SHOULDERS is very bad

- Q. Please indicate the extent to which you are favorable or unfavorable towards HEAD AND SHOULDERS brand hair shampoo.

Most favorable \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:Most unfavorable

- Q. Please indicate the extent to which you like or dislike HEAD AND SHOULDERS.

In general, I like it very much \_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: In general, I don't like it at all

The procedure of gathering attitudinal data (question wording, format, scaling, etc.) were similar to the one utilized by Rosenberg (1956), Fishbein (1963, 1967), and Sheth (1969, 1971, 1974) with one exception: where the 11-point and 21-point scaling used by Rosenberg model were changed to 7-point scales due to problem of categorization in more calibrated scales (Sheth and Park, 1973; Green and Rao, 1970; Miller, 1956).

#### Results and Discussion

The data analysis consisted of regressing attitudes toward Head and Shoulders, Prell and Breck shampoos, as measured by three different ways (good-bad, like-dislike and favorable-unfavorable opinions of the brand) on the cognitive components of each of the three models. In the case of the Rosenberg model, this involved first calculating a weighted sum score for each individual for each brand from  $PI_i$  and  $VI_i$  measures as operationalized by Rosenberg. Then attitudes were regressed on this weighted sum score. The resultant regression of attitude toward the brand on this weighted sum score is by definition a simple regression. In the Fishbein model, a weighted sum score for each individual was calculated for each brand from the profile of  $B_i$  and  $a_i$  scales as suggested by Fishbein. Then attitude toward the brand was regressed on this weighted sum score resulting in a simple regression. Finally, the Sheth model also entailed two steps. The first step consisted of reducing the rank of brand potential profile to its underlying dimensionality with the use of Eckart-Young decomposition procedures. Three underlying dimensions were found consistently across the three brand profiles explaining about 97 percent of the total variance. The rotated dimensions were easy to interpret; the first dimension reflected potential of the shampoo to maintain hair and scalp health (dandruff control, conditioning of hair, and scalp condition); the second dimension reflected potential of the shampoo to provide convenience in shampooing (lather, cleaning of hair, manageable hair); and the third dimension reflected the pleasantness of smell the shampoos provided. The second step entailed regressing attitude toward a brand on the three dimension scores generated in the first step. This involved a multiple regression procedure in which the beta weights estimated for each dimension (hair-scalp health, shampooing convenience and pleasant smell) represented the relative control each type of benefit had in determining a person's attitude toward a brand of shampoo.

The regression of three different attitude measures on the same

cognitive components in each model provides us with the relative efficacy of the models with respect to consistency of relationship. Similarly, the regressions of attitudes for three different brands provides us with an estimate of the generalizability of the models across different attitudinal objects within the same product class. In order to compare the stability of relationship for each model, the total sample for each brand was split into two random subsamples. The change in the regression relationships between the two random subsamples will indicate the stability or robustness of the model when subjected to sampling variations. The final criterion of predictive efficacy entailed building a model in one sample and predicting the attitude scores in another independent sample. We achieved this by using one random subsample to estimate the parameters of the relationship in each model and then predicting attitude scores in the other sample. The correlation between the predicted and the actual attitude score represents the predictive validation of the model. In order to estimate the reliability of this prediction, we utilized the double-cross validation procedure suggested by Tatsuoka (1969) and others. The change in the two predictions gives us an indication of the reliability aspect of predictive efficacies of models.

In order to utilize the four criteria of predictive efficacy of a model (generalizability, consistency, stability and reliability) a total of 81 separate regressions were performed as well as calculating 54 correlations between predicted and actual values of attitudes across three brands and three attitude measures for the three models. In order to remove the upward bias due to two additional variables in the Sheth model and also to remove the upward bias due to the sample size (degrees of freedom), all regression and correlation results were adjusted to make them comparable. The results are summarized in Table 1.

As can be seen from the results, the Sheth model has consistently higher correlations across all the three brands. In fact, the correlations of the Sheth model are all significantly better than the Rosenberg and the Fishbein models. In terms of the generalizability criterion, we can conclude that the Sheth model was the best and the Rosenberg model the poorest with the Fishbein model falling in between the other two models. In general, all the three models do well in the case of Head and Shoulders and Breck but not in the case of Prell shampoo. With respect to the consistency criterion, again the Sheth model provides consistently higher correlations across all three measures of attitudes as compared to the Rosenberg and the Fishbein models. In general, all the three models work best when attitudes are measured on a "good-bad" scale.

With respect to the stability criterion, the average absolute difference between the two randomly split samples across three brands and three measures of attitudes is highest for the Fishbein model (0.18) and lowest for the Sheth model (0.12). Thus, the Sheth model is more stable or robust when subjected to sampling variations. However, surprisingly the Rosenberg model also does as well as the Sheth model (0.12).

TABLE I

Attitude Correlations with the Rosenberg, Fishbein & Sheth Models (Adjusted for Variables and Sample Sizes)

	Head and Shoulders						Prell						Breck														
	Attitude Scale			Attitude Scale			Attitude Scale			Attitude Scale			Attitude Scale			Attitude Scale											
	Good	Favor	Like	Bad	Unfavor	Dislike	Good	Favor	Like	Bad	Unfavor	Dislike	Good	Favor	Like	Bad	Unfavor	Dislike	Good	Favor	Like	Bad	Unfavor	Dislike	Average		
<u>Rosenberg Model</u>																											
1. Total Sample	.57	.53	.50	.52	.42	.47	.52	.42	.47	.57	.62	.58	.53	.49	.49	.55	.55	.55	.55	.55	.55	.55	.55	.55	.55	.53	
2. First Half Estimation	.51	.54	.51	.40	.25	.33	.40	.25	.33	.61	.62	.63	.49	.61	.62	.63	.62	.62	.62	.62	.62	.62	.62	.62	.62	.63	.49
3. Second Half Estimation	.61	.52	.49	.60	.53	.56	.60	.53	.56	.52	.61	.52	.55	.52	.61	.52	.61	.61	.61	.61	.61	.61	.61	.61	.52	.55	
4. Stability Estimation (Difference)	.10	.02	.02	.20	.28	.23	.20	.28	.23	.09	.01	.11	.12	.09	.01	.11	.01	.01	.01	.01	.01	.01	.01	.01	.11	.12	
5. First Half Validation	.51	.54	.51	.40	.25	.33	.40	.25	.33	.61	.62	.63	.49	.61	.62	.63	.62	.62	.62	.62	.62	.62	.62	.62	.63	.49	
6. Second Half Validation	.61	.52	.49	.60	.53	.56	.60	.53	.56	.52	.61	.52	.55	.52	.61	.52	.61	.61	.61	.61	.61	.61	.61	.61	.52	.55	
7. Reliability of Prediction	.09	.01	.01	.20	.28	.23	.20	.28	.23	.09	.00	.11	.12	.09	.00	.11	.00	.00	.00	.00	.00	.00	.00	.00	.11	.12	
<u>Fishbein Model</u>																											
1. Total Sample	.66	.63	.60	.59	.47	.55	.59	.47	.55	.63	.63	.61	.60	.63	.63	.63	.63	.63	.63	.63	.63	.63	.63	.63	.63	.61	.60
2. First Half Estimation	.59	.64	.65	.44	.28	.38	.44	.28	.38	.71	.69	.71	.57	.71	.69	.71	.69	.69	.69	.69	.69	.69	.69	.69	.71	.57	
3. Second Half Estimation	.73	.61	.55	.68	.60	.65	.68	.60	.65	.53	.56	.50	.60	.53	.56	.50	.56	.56	.56	.56	.56	.56	.56	.56	.50	.60	
4. Stability of Estimation (Difference)	.14	.03	.10	.24	.32	.27	.24	.32	.27	.18	.13	.21	.18	.18	.13	.21	.13	.13	.13	.13	.13	.13	.13	.13	.21	.18	

TABLE I  
(Continued)

Attitude Correlations with the Rosenberg, Fishbein  
& Sheth Models (Adjusted for Variables and Sample Sizes)

	Head and Shoulders						Prell						Breck						
	Good		Like		Good		Favor		Like		Good		Favor		Like		Average		
	Bad	Unfavor	Dislike	Dislike	Bad	Bad	Unfavor	Unfavor	Dislike	Dislike	Bad	Bad	Unfavor	Unfavor	Dislike	Dislike	Like	Dislike	Values
5. First Half Validation	.59	.64	.65	.65	.44	.44	.28	.28	.38	.38	.71	.71	.69	.69	.71	.71	.71	.71	.57
6. Second Half Validation	.73	.61	.55	.55	.68	.68	.60	.60	.65	.65	.53	.53	.56	.56	.50	.50	.50	.50	.60
7. Reliability of Prediction (Difference)	.14	.03	.10	.10	.24	.24	.32	.32	.27	.27	.18	.18	.13	.13	.21	.21	.21	.21	.18
<u>Sheth Model</u>																			
1. Total Sample	.73	.73	.72	.72	.64	.64	.52	.52	.60	.60	.78	.78	.76	.76	.73	.73	.73	.73	.69
2. First Half Estimation	.69	.71	.69	.69	.51	.51	.34	.34	.47	.47	.80	.80	.79	.79	.78	.78	.78	.78	.64
3. Second Half Estimation	.76	.75	.75	.75	.73	.73	.63	.63	.70	.70	.77	.77	.75	.75	.67	.67	.67	.67	.72
4. Stability of Estimation (Difference)	.07	.04	.06	.06	.22	.22	.29	.29	.23	.23	.03	.03	.04	.04	.11	.11	.11	.11	.12
5. First Half Validation	.69	.71	.68	.68	.48	.48	.33	.33	.42	.42	.77	.77	.75	.75	.77	.77	.77	.77	.62
6. Second Half Validation	.76	.76	.74	.74	.71	.71	.62	.62	.67	.67	.75	.75	.73	.73	.76	.76	.76	.76	.71
7. Reliability of Prediction (Difference)	.07	.05	.06	.06	.23	.23	.29	.29	.25	.25	.02	.02	.02	.02	.01	.01	.01	.01	.11

Finally, it is again obvious from the results that predictions of attitudes are better when the Sheth model is used compared to the other two models. The average predictive correlation is 0.67 for the Sheth model, 0.58 for the Fishbein model and 0.52 for the Rosenberg model. More importantly, the Sheth model provides more reliable predictions across the two samples. The average absolute difference in predictions is 0.11 in the Sheth model, 0.12 in the Rosenberg model and 0.18 in the Fishbein model.

To summarize, it appears that the Sheth model performs better on all the four criteria of predictive efficacy in relating cognitive beliefs about an object with attitudes toward that object. We believe this is probably due to the following three reasons. First, beliefs are measured only within the realistic and relevant ranges and not on true bipolar scales. The distributions, therefore, tend to be less skewed than in the Rosenberg and the Fishbein models. Second, beliefs are worded in the language of the consumer. The anchorage points in the Sheth model are based on prior pilot study of the consumers making the scales more meaningful to the subjects. We believe this reduces the measurement error which provides consistency in the Sheth model. Finally, the beliefs are not aggregated a priori which removes the assumption of unidimensionality of the cognitive structure underlying attitudes.

Even though the Sheth model performed best of all the three models, it explained on the average only about 46 percent of variance in attitudes toward brands of shampoo. This implies that factors other than cognitive beliefs may also be determining a person's attitude toward a product. Sheth (1974) has suggested that at least three other factors also determine a person's attitude toward a brand in addition to his cognitive beliefs of the type measured in the expectancy-value models. First, a person may like or dislike a product strictly due to past habit which has been reinforced. In other words, the noncognitive process of classical conditioning may generate affective tendencies as suggested by Katz and Stotland (1959). Second, a person may like or dislike a product due to its social imagery which may be independent of the product attributes. Often, the social imagery is in conflict with the brand's potential to satisfy buying motives. Finally, a person may like or dislike a product due to other motivational influences which are not specific or relevant to the product class. Howard and Sheth (1969), for example, have emphasized the important role of curiosity and novelty needs in consumption behavior.

It is clear from this study that consumer behavior is more complex and cannot be fully explained or predicted by the specific expectancy-value models such as the Rosenberg or the Fishbein models. Rather than borrow the models and test to what extent they are relevant in consumer behavior, it seems more advantageous to work backwards by first fully understanding the phenomenon of consumer attitudes and then building indigenous models by judiciously borrowing concepts from relevant disciplines. It is unfortunate that we have not learned our lessons from numerous past futile efforts of blindly borrowing models from the behavioral and the quantitative sciences.

## FOOTNOTES

1. Secil Tuncalp is a Research Assistant Professor, Survey Research Laboratory, and Jagdish N. Sheth is I.B.A. Distinguished Professor and Research Professor at the University of Illinois at Urbana-Champaign.

## REFERENCES

- Alpert, M. I. Identification of determinant attitudes: A comparison of methods. Journal of Marketing Research, 1971, 8, 184-91.
- Banks, S. The relationship between preference and purchase behavior. Journal of Marketing, 1950, 15, 145-57.
- Bass, F. M. & Wilkie, W. L. A comparative analysis of attitudinal predictions of brand preference. Journal of Marketing Research, 1973, 10, 262-269.
- Bem, D. J. Beliefs, attitudes, and human affairs. Belmont, California: Brooks/Cole, 1970.
- Cohen, J. B. & Ahtola, O. T. An expectancy x value analysis of the relationship between consumer attitudes and behavior. In D. M. Gardner (Ed.), Proceedings 2nd Annual Conference Association for Consumer Research, September 1971.
- Cohen, J. B., Fishbein, M. & Ahtola, O. T. The nature and uses of expectancy-value models in consumer attitude research. Journal of Marketing Research, 1972, 9, 456-460.
- Day, G. S. Evaluating models of attitude structure. Journal of Marketing Research, 1972, 9, 279-86.
- Fishbein, M. An investigation of the relationships between beliefs about an object and the attitude toward that object. Human Relations, 1963, 16, 233-239.
- Fishbein, M. (Ed.) Readings in attitude theory and measurement. New York: Wiley, 1967.
- Green, P. E. & Rao, V. R. Rating scales and information recovery-How many scales and response categories to use? Journal of Marketing, 1970, 34 (3), 33-39.
- Hansen, F. Consumer behavior: A cognitive theory. New York: McMillan Company, 1972.
- Howard, J. and Sheth, J. N. The theory of buyer behavior. New York: Wiley, 1969.

- Katz, D. and Stotland, E. A preliminary statement to a theory of attitude structure and change. In S. Koch (Ed.), Psychology: A study of a science. New York: McGraw Hill, 1959.
- McGuire, W. J. The nature of attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), The handbook of social psychology. (2nd Ed.) Massachusetts: Addison-Wesley, 1969.
- Miller, G. A. The magical number seven plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.
- Myers, J. H. and Alpert, M. I. Determinant buying attitudes: Meaning and measurement. Journal of Marketing, 1968, 32 (4), 13-20.
- Raju, P. S. and Sheth, J. N. Nonlinear and noncompensatory relationships in attitudes research. Paper presented at the AMA Educators Conference, Portland, August 11-13, 1974.
- Rosenberg, M. J. Cognitive structure and attitudinal affect. The Journal of Abnormal and Social Psychology, 1956, 53, 367-372.
- Rosenberg, M. J. An analysis of affective-cognitive consistency. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson & J. W. Brehm (Eds.), Attitude organization and change. Yale University Press, 1960.
- Rosenberg, M. J. Mathematical models of social behavior. In G. Lindzey & E. Aronson (Eds.), The handbook of social psychology, Vol. 1, Massachusetts: Addison-Wesley, 1968.
- Sheth, J. N. Attitude as a function of evaluative beliefs. Paper presented at AMA Consumer Workshop, Columbus, Ohio, August 1969.
- Sheth, J. N. Affect, behavioral intention, and buying behavior as a function of evaluative beliefs. In P. Pellmans (Ed.), Insights in consumer and market behavior. Belgium: Namur University Publications, 1971.
- Sheth, J. N. Reply to comments on the nature and uses of expectancy-value models in consumer attitude research. Journal of Marketing Research, 1972, 9, 462-465.
- Sheth, J. N. & Talarzyk, W. W. Perceived instrumentality and value importance as determinants of attitudes. Journal of Marketing Research, 1972, 9, 6-9.
- Sheth, J. N. Brand profiles from beliefs and importances. Journal of Advertising Research, 1973, 13 (1), 37-42.



- Sheth, J. N. & Park, C. W. Equivalence of Fishbein and Rosenberg theories of attitudes. Proceedings of the 1973 American Psychological Association Conference, Montreal, 1973.
- Sheth, J. N. A field study of attitude structure and attitude-behavior relationship. In J. N. Sheth (Ed.), Models of buyer behavior. New York: Harper & Row, 1974.
- Sheth, J. N. The next decade of buyer behavior theory and research. J. N. Sheth (Ed.), Models of buyer behavior. New York: Harper & Row, 1974.
- Tatsuoka, M. M. Validation studies: The use of multiple-regression equations. Champaign, Illinois: Institute for Personality and Ability Testing, 1969.
- Thurstone, L. L. Attitudes can be measured. American Journal of Sociology, 1928, 33, 529-554.
- Triandis, H. C. Attitude and attitude change. New York: Wiley, 1971.
- Wilkie, W. L. and Pessemier, E. A. Issues in marketing's use of multi-attribute models. Journal of Marketing Research, 1973, 10, 428-441.

PREDICTIVE VALIDATION AND CROSS-VALIDATION OF  
THE FISHBEIN, ROSENBERG, AND SHETH MODELS  
OF ATTITUDES

P.S. Raju, Rabi S. Bhagat, and Jagdish N. Sheth<sup>1</sup>  
University of Illinois

The present study compares the Fishbein, the Rosenberg and the Sheth models in their ability to predict behavioral intentions of consumers with respect to the 'Pinto' car. Since the Rosenberg model is restricted to the prediction of 'attitudinal affect' it has been extended to the prediction of behavioral intention for the purposes of this study. Three stages have been identified in the comparison process namely predictive validation, cross-validation and validity generalization. Predictive validation deals with the comparison based on one sample, cross-validation extends the comparisons to other samples from the same population, and validity generalization indicates the extent to which the results are applicable to other populations. While most studies have been restricted to the stage of predictive validation, the present study also includes cross-validation. Data were obtained from 243 respondents, comprising of students and housewives, in the Champaign-Urbana area. The results indicate that the Sheth model has high predictive validity and cross-validity, the Fishbein model has a lower predictive validity but high cross-validity, and the Rosenberg model has a low cross-validity so that not much confidence can be placed in its predictive validity over different samples.

Although the concept of attitude has been researched for almost a century, the development of comprehensive attitudinal models is a relatively recent phenomenon. The recent efforts devoted to the development of these models can be attributed to the following reasons: 1) they provide a systematic framework for research by specifying variables that relate to the process, 2) they aid in the generation of new hypotheses, and 3) they provide diagnostic information which facilitates the determination of future courses of action.

Based on cognitive consistency theories, Rosenberg (1956) showed that attitudinal affect could be expressed as a simple mathematical function of cognitive elements. After his pioneering work, several researchers in the areas of social psychology and consumer psychology have developed other models of attitudes. However, the emphasis seems to have shifted from predicting affective tendencies toward the object to the prediction of behavioral intention and specific behaviors toward the object. For instance, Fishbein and his

associates (Ajzen & Fishbein, 1970(a), 1970(b), 1973; Fishbein, 1967, 1972) have formulated a model in which behavioral intention is predicted from personal beliefs as well as social normative variables. The Rosenberg and the Fishbein models have been very instrumental in fostering the expectancy value tradition in attitudinal research. In the area of consumer behavior, for example, Wilkie and Pessemier (1973) report that, apart from numerous studies conducted in Europe, over forty studies have been published in the United States alone within a period of three years. Unfortunately, there is little, if any, consistency among these studies in terms of the operationalization or measurement of constructs, so that their usefulness in drawing broad generalizations is rather limited. In addition, expectancy-value models have often been criticized for their inherent limitations (Day, 1972; Sheth, 1973; Sheth & Tuncalp, 1974; Wilkie & Pessemier, 1973). Finally, a number of researchers in social psychology have conceded that attitudes toward the object, no matter how they are defined and measured are not the sole determinants of a person's behavioral intention or actual behavior toward the object (Katz & Stotland, 1959; Rokeach, 1968; Triandis, 1971). Similar suggestions have also been made in Consumer Psychology and in Marketing (Howard & Sheth, 1969; Sandell, 1968; Sheth, 1971; Sheth & Raju, 1973). Recently, Sheth (1971, 1974) has proposed a model of brand choice behavior which attempts to integrate diverse thinking on the role of attitudes in the prediction of behavior both in social psychology and in consumer psychology. The model states that behavioral intentions and behavior of an individual are a function of 1) evaluative beliefs about the object's potential to satisfy his needs, wants, or desires, 2) social stereotype of the object as perceived by him, 3) past satisfaction and resultant predisposition (affect) toward the object, and 4) situational events which are either anticipated or which may unexpectedly occur at the time of behavior.

#### Criteria Used to Compare Models

In spite of the availability of a variety of models, little has been done in terms of comparative research on these models. As a result, the researcher is often unclear as to the relative strengths and weaknesses of different models, and their appropriateness for different situations. "Unless we do a comparative study under the same setting, on the same issue, and on the same or essentially the same subjects, it is not possible to claim victory for any model" (Sheth, 1972, p. 465). Comparative studies, therefore, need to be conducted over a variety of situations and behaviors before conclusive assertions can be made about the appropriateness of any particular model for a problem area.

Researchers have often utilized one or more of the following four major criteria in comparing competing models or theories:

- 1) Description: this indicates the ability of the model to describe the process or phenomenon being studied. The use of flow diagrams can greatly aid the descriptive capacity of a model.
- 2) Explanation: the extent to which a model can provide the causes from the knowledge of the effects determines its rating on this criterion.

- 3) Prediction: the capability of a model to predict future occurrences based on the knowledge of certain key concepts is indicated by the rating on this criterion. While explanatory models are usually predictive, the reverse need not necessarily be true (Kaplan, 1964). For instance stochastic models of brand choice are predictive but do not explain the process.
- 4) Prescription: Often referred to as also the diagnostic ability of a model it essentially deals with the capacity of the model to prescribe courses of action to influence the process in a desired manner.

In conformity with the views presented above, the main aim of this study is to compare the Rosenberg, the Fishbein, and the Sheth models in their ability to predict behavioral intentions of consumers with respect to a brand of a durable consumer product. We restrict ourselves to the single criterion of 'prediction' due to two major reasons:

- 1) It is beyond the scope of this study to attempt a comparison on all the relevant criteria, and
- 2) The criteria of 'prediction' and 'prescription' are considered most important and have been used frequently in past research to evaluate attitude models. A comparison on the prescriptive criterion would, however, require a longitudinal study since we would be concerned with the implications of the model for future action. As this study is cross-sectional it is infeasible to test the prescriptive efficacy of the models.

An exhaustive comparison of the models on the basis of 'prediction' would require a three stage analysis comprising of predictive validation, cross-validation, and validity generalization. Since all the models use a multiple regression formulation, as will be discussed later, the three stages are described below with respect to that statistical technique.

Predictive validation provides an indication of the extent to which the model is able to predict the criterion variable for a particular sample representative of the population. A model is considered valid if the correlation coefficient and, therefore, the percentage variance of the criterion variable explained is quite large. By predictive validation alone, however, it is impossible to generalize the results even to other samples from the same population unless the sample size is so large as to cover the entire population. 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. Mosier (1951) in an excellent paper on cross-validation states:

"Since both combining weights ( $\beta$  weights) and their effectiveness should be determined on samples representative of the group in which the battery (in our case 'model') will be applied, it is obvious that the two samples should be selected from the same universe." (p.6)

In cross-validation, a single sample is divided randomly into a 'derivation sample' and a 'validation sample' or two independent random samples are drawn from the population. The multiple regression weights and the multiple correlation coefficient are obtained by applying the model to the derivation sample. The obtained regression weights are used in the validation sample to create 'predicted criterion scores' which are then correlated with the

actual criterion measures for the validation sample. The extent to which this correlation coefficient is close to that obtained initially in the derivation sample determines the cross-validity of the model. For perfect cross-validity the correlation coefficients should be equal. If a model has good cross-validity it can be presumed that the predictive power will be almost constant over different samples from the same population.

The test for validity generalization is very similar to that for cross-validity except that the multiple regression weights and the multiple correlation coefficient obtained on a sample drawn from one population are tested on a second sample drawn from a different population. The extent of validity therefore signifies the applicability of a model to different populations.

With respect to the typology of validation presented above, it is interesting to note that most validation studies on attitude models have been restricted to the stage of 'predictive validation' thereby providing biased results, with no indication of the generalizability to other samples or other populations. In this study, the validation procedure has been designed to include cross-validation in addition to predictive validation. Validity generalization has, however, not been performed because our interest at this stage is limited to samples from a single population. In summary, this study was designed to accomplish three objectives:

- 1) A detailed comparison of the Rosenberg, the Fishbein, and the Sheth models, both conceptually and in terms of their ability to predict buying intentions of consumers.
- 2) Predictive validation of the models based on results obtained from the total (derivation + validation) sample.
- 3) Cross-validation of the models by applying the model coefficients obtained from the derivation sample to the validation sample.

#### Brief Descriptions of Models

A brief description of each of the three models used in the study is provided. A more elaborate treatment of each model is available in the previous writings of Fishbein (1967, 1972), Rosenberg (1956), and Sheth (1974).

#### The Rosenberg Model

As a means to understanding the attitude learning and attitude change processes, Rosenberg (1956, 1960) formulated a structural relationship between the attitude toward an object and the beliefs about the object. The relative stable affective orientation [ $A_o$ ] toward an object or concept is considered to be a function of two cognitive variables: 1) Value Importance [VI]: the degree of satisfaction generated by the attainment of a desired value state; and 2) Perceived Instrumentality [PI]: the extent to which the object or concept leads to or blocks the attainment of the value state. Every individual, therefore, develops an attitude in conformity with his basic system of values. Mathematically, the model is stated as

$$A_o = \sum_{i=1}^n (PI_i) (VI_i)$$

where  $n$  is the number of desired value states.

Since Rosenberg was mainly concerned with the cognitive processes underlying attitude formation and change, he has restricted his model to the prediction of attitude. However, our interest in this paper centers on the prediction of behavioral intention. Rosenberg (1960) has indicated, in the context of action predicting capability of attitudes, that although attitudes are not the sole predictors of behavior, they could be effective predictors. He also recognizes the presence of situational factors and restraints that could possibly influence behavior. Since Rosenberg has not extended his model to include these situational factors, we shall restrict ourselves to the prediction of behavioral intention from the attitude score obtained from his model. The Rosenberg model, thus consists of a single determinant construct namely the personal beliefs about the role of the object in attaining differentially valued states or goals.

#### The Fishbein Model

According to Fishbein (1967, 1972), behavioral intention is a function of two components; one attitudinal and the other normative. The model is represented formally as

$$B \sim BI = [A_{act}] w_0 + [(NB)(Mc)] w_1$$

where B = behavior, BI = behavioral intention,  $A_{act}$  = attitude toward the act or behavior, NB = normative belief, i.e., what others expect or say should be done in the situation, Mc = motivation to comply with the normative belief, and  $w_0$  and  $w_1$  are empirically determined weights.  $A_{act}$  itself is determined as

$$\sum_{i=1}^n B_i a_i$$

where  $B_i$  = the individual's belief about the likelihood that the behavior will result in the  $i$ th consequence or outcome,  $a_i$  = the individual's evaluation of outcome  $i$ , and  $n$  = number of salient beliefs.

There are three aspects of the model which need to be pointed out. First, behavioral intention is considered to be an immediate antecedent of behavior so that other influences that might be active between the two are ignored. In the context of consumer behavior this assumption may not be valid because of the considerable time span that mediates the expression of an intention and its translation into action (Howard and Sheth, 1969; Sheth, 1971). Second, the model explicitly recognizes the importance of variables such as social norms and motivation in influencing behavior. In this sense it is a definite improvement over the Rosenberg model. However, all other variables influencing behavior are presumed to operate indirectly by influencing the two components in the model or their relative weights. Finally the model makes a clear distinction between 'attitude toward the object' and 'attitude toward the act'; only the latter being considered an effective predictor of behavior.

#### The Sheth Model

In the comprehensive model proposed by Sheth (1971, 1974) behavior [B] is determined by 1) the behavioral intention with respect to the object [BI], 2) past satisfaction from behavior which results in a predisposition toward the object [S], and 3) Unexpected events [UE] that influence the

individual at the time of manifestation of the behavior. Mathematically,

$$B = w_1 [BI] + w_2 [S] + w_3 [UE].$$

The unexpected events factor is presumed to be uncorrelated with the predisposition or the behavioral intention and it may either enhance or inhibit the translation of intentions to actions. It is expected that, as the time period between intention and behavior increases, the influence of this factor will correspondingly increase; for example, unexpected events play a major role in the purchase of consumer durable goods.

Although the Sheth model can be used to predict behavior, our interest in this study is limited to behavioral intention, which is hypothesized to be a function of 1) evaluative beliefs [EB] about the object's potential to satisfy needs, wants, and desires, 2) perceived social stereotype [SS] of the object, 3) predisposition [S] resulting from past satisfaction, and 4) situational influences [AS] that the person anticipates will be effective at the time of behavior. In a functional form,

$$BI = w_1 [EB] + w_2 [SS] + w_3 [S] + w_4 [AS].$$

Evaluative beliefs represent the individual's perception of the object's capability to satisfy a set of relevant needs, wants and desires. In this respect, evaluative beliefs closely resemble the perceived instrumentality concept in the Rosenberg model, the major difference being in the definition of valued states. While Rosenberg model has been limited to more fundamental values of the individual, the valued states in the Sheth model are specific to the class of objects on which research is conducted. Evaluative beliefs are directly related to the instrumental-utilitarian function that attitudes are presumed to perform in Katz's functional formulation (1960). However, evaluative beliefs are presumed to be multidimensional which requires a profile analysis of the object's capability to satisfy diverse functional needs of the individual.

Social stereotype represents the individual's perceptions of the social connotations or social imagery the object possesses. For example, a brand of cigarettes such as Virginia Slims may be perceived to be more feminine than other brands such as Marlboro although functionally the two cigarettes are highly similar in their characteristics. The social stereotype is presumed to be due to the identification of the object with specific segments of the society based on life cycle factors (age, marital status, children, etc.), on socioeconomic factors (education, occupation, income, housing, mobility, etc.), on life style factors (activities, interests, opinions, and values, etc.), and cultural factors (religion, tradition, language, etc.) and biological factors (sex, race, etc.). The Sheth model presumes that social stereotype is also a multidimensional concept which requires dimensional analysis on a profile of perceptions of the individual about the goal object as a social object.

Anticipated situation pertains to all the relevant situational influences that the individual expects will occur, and all the relevant activities that the individual expects he will engage in at or around the time of manifestation of behavior. Expectation of certain occurrences would therefore enhance the intentions whereas expectation of certain others might inhibit them. Events such as vacations, moving, and marriage are usually anticipated by a person in determining his intentions.

Finally, Satisfaction represents the positive or negative predisposition toward an object based on past experience with the object as instrumental to the attainment of certain desirable goals. Satisfaction is measured by the individual's affective tendency toward the object.

For a more elaborate treatment of the constructs the reader is referred to the earlier writings of Sheth (1971, 1974).

### A Conceptual Comparison of the Three Models

In this section, we will briefly point out similarities and differences among the models with respect to three broad areas; number of constructs, definition and measurement of constructs and operationalization of the models.

#### Number of Constructs

The extended Rosenberg model used in this study is the simplest of the three models since it uses only one construct namely 'attitudinal affect' in the prediction of behavioral intention. While recognizing the importance of the attitudinal variable, Fishbein also incorporates in his model social normative influences and hence takes into account the fact that behavior is not a function of attitude alone. While definitely an improvement over the Rosenberg model, the Fishbein model still does not consider the effect of other important variables such as situational influences. The 'evaluative belief' and the 'social stereotype' constructs in the Sheth model (restricting to the behavioral intention level) are roughly the counterparts of the 'attitudinal' and 'social normative' constructs respectively of the Fishbein model. Sheth, however, has two additional constructs that make the model more comprehensive. One of these constructs 'predisposition resulting from past satisfaction' takes into account the fact that affect need not always be a result of the beliefs held by the individual. The other construct 'Anticipated situations' recognizes the importance of situational influences that a person is able to foresee in determining his intentions.

Finally, in contrast to the assumption made by Fishbein that behavioral intention is an immediate antecedent of behavior Sheth explicitly considers unexpected events to intervene between the two. However, research still needs to be done to effectively operationalize this construct. Recent research efforts concerning situational influences on behavior (Belk, 1974; Sandell, 1968) have provided valuable insight into the problem.

#### Definition and Measurement of Constructs

The models can be compared with respect to three distinct types of constructs that are utilized in the prediction of behavioral intention: 1) Attitudinal 2) Social and 3) Situational. In addition, there are important differences between the Fishbein and the Sheth models in the measurement of behavioral intention. All three models use the attitude type construct, only the Fishbein and the Sheth models use the social type construct, and the situational type construct is used only by the Sheth model. As a result, no comparison can be made with respect to the measurement of situational constructs, the comparison is restricted to the Fishbein and the Sheth models for the social type constructs and all three models are compared with respect to the attitudinal construct.



All the three models consider 'attitude' to be a univariate unidimensional construct; but measure it in different ways. In the Rosenberg model it is a summated product of perceived instrumentality and value importance. Rosenberg recommends a 21 point scale ranging from -11 to +11 to measure value importance and a 11 point scale ranging from -5 to +5 to measure perceived instrumentality. It has been suggested, especially after the work of Miller (1956), that such scales might be inappropriate because a person usually cannot distinguish more than five to nine different levels of a concept (Sheth & Park, 1973; Tuncalp, 1973). In this study, therefore, these components have only been measured using scales ranging from -3 to +3, although the wording of the scales is consistent with the recommendations of Rosenberg. Fishbein also uses the expectancy-value approach and the two components of 'belief' and 'evaluation' are measured by scales ranging from -3 to +3. However, Fishbein measures 'attitude toward the act' as opposed to 'attitude toward the object'. In the Sheth model, the evaluative beliefs are not combined into a single construct called 'attitude', but there are essentially the counterpart of such a construct as used by Rosenberg and Fishbein. The evaluative beliefs are measured by scales ranging from 1 to 7 and a further distinction is that all the scales need not necessarily be bipolar. They are anchored at the ends by the extremities of the normal range of the belief that a person uses. For example, the scale for the evaluative belief regarding durability of Pinto ranges from 'very durable' to 'only average in durability' rather than from 'very durable' to 'not at all durable'. Further differences in the wording of the scales for the three models will become evident from the sample of scales that will be presented in the 'Method' section.

The 'social stereotype' construct considered by Sheth is quite different from the social normative construct as defined by Fishbein. While Fishbein is mainly concerned with what others want the person to do and whether the person wants to comply with their wishes, Sheth is concerned with the social connotations and social imagery of the object. Again Fishbein uses scales ranging from -3 to +3 to measure each component of the social normative construct whereas Sheth uses scales ranging from 1 to 7 to measure social stereotypes.

Finally, a major distinction between the Fishbein and the Sheth models is in the measurement of behavioral intention, Fishbein has convincingly argued that for better prediction, behavioral intention should be measured with respect to a specific object rather than a generalized group of objects (Fishbein, 1967). Thus, buying intention should be measured with respect to 'Pinto' rather than toward the general product category of 'automobile'. While Sheth agrees with this point, he measures buying intention not only with respect to a specific brand such as 'Pinto' but also qualifies it with respect to the need or motivational level of the buyer. His scale, therefore, measures how seriously a person would consider buying a 'Pinto' if he were to buy an automobile. Since Rosenberg does not provide a scale for measurement of behavioral intention both the Fishbein and the Sheth scales have been used as criterion measures in the Rosenberg model. The Sheth scale, however, has been recoded from -3 to +3 (instead of 1 to 7) when used on the Rosenberg model to make it compatible with the way 'perceived instrumentality' and 'value importance' are measured.

### Operationalization of Models

The major difference in the operationalization of the models is the expectancy-value approach adopted by Rosenberg and Fishbein versus the factor analytic approach adopted by Sheth. In the expectancy-value approach, the attitude is determined by the summation of the two components of expectancy and value over the whole range of values (or expectancies). In the process, the positive and negative components cancel each other out leaving a simplified index of the cognitive structure. Such an assumption may not be true. Researchers have suggested that a disaggregated version of the model might perform better than the summated version (Gohen & Ahtola, 1971; Lutz & Howard, 1972; Sheth, 1970). In contrast to this, Sheth performs a dimensional analysis on the profile of evaluative beliefs by the technique of principal components analysis. The resultant orthogonal dimensions of the evaluative beliefs are then utilized as independent predictors of behavioral intention. Such a factor analytic approach has several advantages:

First, it is a compromise between the aggregated version and the completely disaggregate version. The disadvantage of the former has already been pointed out. In the latter, all the beliefs are used as independent predictors of behavioral intention. This could result in too many independent variables especially in cases where there are a number of salient beliefs e.g., automobile purchase. By utilizing only the orthogonal dimensions of the beliefs, the number of predictor variables can be reduced considerably without sacrificing too much predictive power. Second, the factor analytic procedure takes into account the multicollinearity among belief items. Thus, in the aggregated version the correlation between belief items leads to double counting of certain dimensions of the beliefs and in the completely disaggregated version it leads to unstable regression weights. These problems are not encountered in Sheth's approach. Finally, the factor analytic approach is very useful from a policy viewpoint. By using attitude models, the marketer is able to gain some insight as to which beliefs will have to be changed to create a desired change in attitude. However, it is often impossible to influence one belief alone without any impact on other cognitive elements. To avoid undesirable consequences on the beliefs that do not need to be influenced, it is necessary for the marketer to understand the dimensions of the cognitive structure and the beliefs that group together under any given dimension. Different promotional and marketing strategies can then be assessed in terms of their effects on the total cognitive structure as opposed to only certain specific beliefs.

#### Method

The models were empirically tested by collecting data on 243 respondents in the Champaign-Urbana area. The respondents were mainly students, student wives, and housewives. An indepth interview with about twenty-five housewives and twenty-five students indicated that the Pinto car (a product of Ford Motor Company) was a salient attitude object for this group of respondents. A separate pilot study involving 40 respondents was performed to obtain salient beliefs pertaining to the purchase of an automobile. Twelve belief items mentioned most frequently were finally selected for the main study. Similar belief items have been utilized as product attributes in past studies (Alpert, 1971). A sample of the scales used for measurement of the various constructs is provided:

Rosenberg Model

1) Value importance (VI)

Please check each scale below so as to indicate the extent to which each consequence that is associated with buying an automobile gives you satisfaction or dissatisfaction.

Economy of operation:

Gives me maximum satisfaction :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: Gives me maximum dissatisfaction  
 (11 other value importances measured similarly).

2) Perceived instrumentality (PI)

Please check each scale so as to indicate to what extent a specific consequence is attained or blocked by buying a Pinto.

Economy of operation:

Completely blocked :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: Completely attained.

Fishbein Model<sup>2</sup>

1) Belief (B<sub>i</sub>)

My buying a Pinto would mean buying an automobile that is economical to operate.

Probable :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: Improbable

(eleven other beliefs were measured similarly).

2) Evaluation of belief (a<sub>i</sub>)

Buying a car that is economical to operate is

Good :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: Bad

(eleven other evaluations were measured similarly).

3) Normative belief (NB)

Others who are important to me think

I should :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_: I should not

buy a Pinto.

4) Motivation to Comply (Mc)

In general, I want to do

:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:

In general, I don't want to do

what others who are important to me think I should do.

5) Behavioral Intention (BI)

I would :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_ I would not  
buy a Pinto.

Sheth Model1) Evaluative Beliefs (EB)

Pinto is economical Pinto is expensive  
to operate :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_ to operate

(eleven other evaluative beliefs were measured similarly).

2) Social Stereotype (SS)

Pinto is meant for young people only.

Strongly agree :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_ Strongly disagree

(eight other Social Stereotypes were measured similarly).

3) Predisposition or Past Satisfaction (S)

Please indicate the extent to which you are favorably or unfavorably  
disposed toward the Pinto

Most favorable toward Most unfavorable toward  
Pinto :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_ Pinto

4) Anticipated Situation (AS)

Three different scales were used to measure personal (AS<sub>1</sub>), buying (AS<sub>2</sub>),  
and financial (AS<sub>3</sub>) situations relating to the purchase of a car.

Personal (AS<sub>1</sub>): Is it conceivable that you might change your intention  
to buy or not to buy an automobile because of some unforeseen events  
(for example, moving, getting married, birth in the family, unanticipated  
change in your financial status or deciding to take vacations) you did  
not anticipate may occur in the next six months?

Not at all conceivable :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_ Very much conceivable

Buying (AS<sub>2</sub>): Is it conceivable that you might change your intentions to  
consider or not to consider buying a Pinto due to some unforeseen events  
including a good buy on some other car or friend's advice to reconsider  
your plans and intentions?

Not at all conceivable :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_ Very much conceivable

Financial (AS<sub>3</sub>): If you were to buy an automobile in the next six months,  
would you have any financial problems?

I simply cannot I can easily raise money  
afford it :\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_:\_\_ to pay for it

5) Behavioral intention (BI)

If you were to buy an automobile, how seriously would you consider buying a Pinto?

Definitely would consider buying a Pinto : \_:\_:\_:\_:\_:\_:\_:\_ : Definitely would not consider buying a Pinto

Since the Sheth model utilizes factors scores for both evaluative beliefs and social stereotypes, principal components analyses were first performed independently on the respective correlation matrices using the total sample. The analyses yielded three major factors in each case and rotated factor loadings were obtained by the technique of orthogonal varimax rotation. Factor scores for each individual were obtained from the rotated factor loadings matrix and these scores were kept invariant for all further analyses. Thus, even when the total sample was split into two for purposes of cross-validation, each individual retained the same factor scores. It is essential that the factor scores be invariant because, in cross-validation, we are concerned with the stability of regression weights which will be confounded by the change in factor scores themselves if these are calculated separately for the derivation and the validation samples.

The predictive validity of each model was determined by a multiple regression technique using behavioral intention as the criterion and appropriate variables as predictors. In the case of the Rosenberg model alone this reduces to simple regression analysis since there is only one predictor variable. For purposes of clarity, the regression equations are reproduced below:

- 1) Rosenberg's model using Fishbein's behavioral intention scale.

$$BI = p_1 \left[ \sum_{i=1}^n (PI_i) (VI_i) \right] + K_1$$

- 2) Rosenberg's model using Sheth's behavioral intention scale.

$$BI = p_2 \left[ \sum_{i=1}^n (PI_i) (VI_i) \right] + K_2$$

- 3) Fishbein model:

$$BI = w_0 \left[ \sum_{i=1}^n B_i a_i \right] + w_1 [NB \times Mc] + K_3$$

- 4) Sheth model:

$$BI = b_1 [EB_1] + b_2 [EB_2] + b_3 [EB_3] + b_4 [SS_1] + b_5 [SS_2] + b_6 [SS_3] \\ + b_7 [S] + b_8 [AS_1] + b_9 [AS_2] + b_{10} [AS_3] + K_4$$

where  $p_1, p_2; w_0, w_1;$  and  $b_1, b_2 \dots b_{10}$  are regression weights and  $K_1, K_2, K_3$  and  $K_4$  are constant intercept values, all determined empirically. In the

Fishbein model  $A_{act}$  was determined by a summated product of  $B_i$  and  $a_i$ . In the Sheth model, it can be noticed that there are three evaluative belief factors ( $EB_1$ ,  $EB_2$ , and  $EB_3$ ) and three social stereotype factors ( $SS_1$ ,  $SS_2$ , and  $SS_3$ ).

The cross-validity of the models was finally determined by utilizing the regression weights obtained from the derivation sample to predict the criterion values in the validation sample. The predicted criterion values were then correlated with the actual criterion values to determine a cross-validation correlation coefficient.

### Results

The results of predictive validation for the Rosenberg model are summarized in Table 1. It can be seen that attitude is a significant predictor of behavioral intention explaining a minimum of about 15 percent of variance in the total sample using Fishbein's behavioral intention scale and a maximum of about 29 percent of the variance in the derivation sample using Sheth's behavioral intention scale. The cross-validation results of the Rosenberg model are not very impressive. The cross-validity correlation coefficient with Fishbein's behavioral intention scale was about 0.23 as opposed to 0.51 obtained in the derivation sample and with Sheth's behavioral intention scale it was about 0.32 as opposed to 0.54 obtained in the derivation sample. Hence we cannot presume with confidence that the performance of the Rosenberg model will be consistent over different samples from the same population.

TABLE 1

Rosenberg Model - Predictive Validation Results

Predictor	Fishbein's BI scale				Sheth's BI scale			
	Total Sample N=243		Derivation Sample N=124		Total Sample N=243		Derivation Sample N=124	
	$\beta$ wt.	Std. Error	$\beta$ wt.	Std. Error	$\beta$ wt.	Std. Error	$\beta$ wt.	Std. Error
$A_0 = \Sigma(PI)(VI)$	0.386	0.059	0.509	0.078	0.438	0.058	0.537	0.076
Multiple correlation (R)	0.386***		0.509***		0.438***		0.537***	
Variance explained ( $R^2$ )	0.149		0.259		0.192		0.288	
Adjusted $R^2$	0.149		0.259		0.192		0.288	
F ratio	42.269		42.709		57.201		49.445	
Std. error of estimate	1.664		1.584		1.781		1.640	

\*\*\*  $p < 0.001$

The predictive validation results for the Fishbein model are presented in Table 2. One interesting finding is that of the two components in the model, 'attitude toward the act' is a significant predictor of behavioral intentions whereas the social normative component is not. Both the predictors,

TABLE 2  
Fishbein Model - Predictive Validation Results

Predictors	Total Sample N=243		Derivation Sample N=124	
	$\beta$ wt.	Std. error	$\beta$ wt.	Std. error
$A_{act} = \sum B_i a_i$	0.467***	0.057	0.485***	0.079
(NBxMc)	-0.037	0.057	-0.043	0.079
Multiple Correlation (R)	0.472***		0.487***	
Variance explained ( $R^2$ )	0.223		0.237	
Adjusted $R^2$	0.220		0.231	
F ratio	34.318		18.823	
Std. error of estimate	1.594		1.614	

\* $p < 0.05$   
\*\* $p < 0.01$   
\*\*\* $p < 0.001$

however, account for about 22 percent variance in behavioral intention in the total sample and about 24 percent in the derivation sample. The predictive ability is, thus, quite stable. Further confirmation of this fact is obtained when the cross-validation correlation coefficient is found to be about 0.432, which is not a substantial drop from the value of 0.487 obtained in the derivation sample. We could, therefore, expect the Fishbein model to consistently predict about 23 percent variance in behavioral intention in other samples drawn from the same population.

Table 3 shows the rotated factor structure of evaluative beliefs obtained for the Sheth model on the total sample. Three major factors were extracted explaining about sixty percent of the total variance. Based on the loadings the factors can be interpreted as follows:

The first factor with high loadings on items such as durability, handling, safety, ride, acceleration, and resale value pertains to an overall sense of the 'quality' dimension of Pinto. Beliefs related to luxury, size of engine,

TABLE 3

Sheth model - Rotated Factor Structure  
of Evaluative Beliefs on Total Sample  
(N=243)

Items	Factor I	Factor II	Factor III	$h^2$
1. Luxury/Economy	-0.077	0.838	0.160	0.733
2. Big/Small Engine	0.119	0.734	0.260	0.621
3. Pollution	0.064	0.723	-0.066	0.531
4. Sportyness	0.035	0.176	0.847	0.749
5. Expensive/Economical to buy	0.118	0.680	0.018	0.476
6. Economical to operate	0.100	-0.608	0.400	0.538
7. Durability	0.731	0.100	-0.103	0.555
8. Good/Poor handling	0.605	-0.179	0.484	0.632
9. Safety	0.827	0.071	0.037	0.691
10. Ride	0.828	0.047	0.048	0.691
11. Accleration	0.648	0.191	0.216	0.503
12. Resale Value	0.689	-0.097	0.038	0.486

Sum of  $h^2$  = 7.207  
Total variance explained = 60.058%

pollution, buying cost, and economy of operation load on the second factor which can be interpreted as a 'luxury' dimension of Pinto. The last factor comprising of beliefs pertaining to sportyness, economy of operation, and handling represents the 'sportyness' dimension of Pinto. The rotated factor structure of social stereotypes presented in Table 4 can also be interpreted in a similar manner. The first factor stereotypes Pinto as a car for 'young unmarried people', the second as a car for 'young people with moderate income', and the third as a car for 'older people with low income'. The three factors together account for 56 percent of the total variance in the social stereotypes.



TABLE 4

Sheth Model - Rotated Factor Structure  
of Social Stereotypes on Total Sample  
(N=243)

Items	Factor I	Factor II	Factor III	h <sup>2</sup>
1. PINTO is meant for young people only	0.146	0.763	0.146	0.625
2. PINTO is meant for people with moderate income	0.118	0.615	0.095	0.401
3. PINTO is suitable for older people	0.092	-0.657	0.095	0.450
4. PINTO is a car meant for everybody	0.180	-0.698	-0.106	0.531
5. PINTO is great as a second car in the family	0.651	-0.084	0.102	0.442
6. Teenagers and College students love PINTO	0.747	0.080	-0.220	0.613
7. Very rich people would never consider buying a PINTO	0.053	0.178	0.867	0.786
8. PINTO is great for a bachelor	0.568	0.012	-0.484	0.557
9. Young unmarried women prefer PINTO	0.784	0.014	0.113	0.628
Sum of h <sup>2</sup>		= 5.032		
Total variance explained		= 55.907%		

Table 5 summarizes the predictive validation results for the Sheth model. There are ten predictor variables in the regression equation; three evaluative belief factors, three social stereotype factors, prior predisposition, and three anticipated situational variables. The three anticipated situational variables pertain to anticipated personal situations (such as moving, marriage, etc.), anticipated buying situation (such as expecting a better buy), and anticipated financial situation. The multiple correlation coefficient for the total sample is about 0.73 and for the derivation sample is about 0.75, explaining about fifty-three percent and about fifty-six percent respectively of the variance in behavioral intention. The four significant predictors of behavioral intention are prior predisposition, anticipated buying situation, the 'quality' dimension of evaluative beliefs, and the 'sportyness' dimension of evaluative beliefs.

TABLE 5  
Sheth Model - Predictive Validation Results

Predictors	Total sample (N=243)		Derivation sample (N=124)	
	$\beta$ Wt.	Std. error	$\beta$ Wt.	Std. error
Evaluative Belief (Factor I)	0.121*	0.054	0.267**	0.080
Evaluative Belief (Factor II)	-0.016	0.046	0.042	0.063
Evaluative Belief (Factor III)	0.128**	0.047	0.170*	0.069
Social Stereotype (Factor I)	-0.005	0.051	-0.035	0.072
Social Stereotype (Factor II)	-0.017	0.047	0.032	0.065
Social Stereotype (Factor III)	-0.024	0.047	-0.051	0.071
Prior Predisposition	0.483***	0.058	0.386***	0.088
Anticipated Situation (Personal)	0.039	0.047	0.037	0.068
Anticipated Situation (Buying)	-0.245***	0.051	-0.239***	0.070
Anticipated Situation (Financial)	-0.049	0.046	-0.069	0.064
Multiple Correlation (R)	0.728***		0.749***	
Variance explained (R <sup>2</sup> )	0.530		0.561	
Adjusted R <sup>2</sup>	0.512		0.526	
F ratio	26.148		14.455	
Std. error of estimate	1.384		1.338	

\* p < 0.05  
 \*\* p < 0.01  
 \*\*\* p < 0.001

The cross-validation correlation coefficient obtained was about 0.665 pointing to the fact that the high correlation coefficients are not due to idiosyncracies of the sample and that the model can be expected to perform equally well in other samples from the same population.

Since the three models use different numbers of predictor variables, the coefficient of determination ( $R^2$ ) cannot be compared directly. Adjusted  $R^2$  values were, therefore, calculated in each case to account for the number of predictor variables and degrees of freedom. The formula used was

$$\text{Adjusted } R^2 = 1 - (1 - R^2) \frac{N-1}{N-n}$$

where  $N$ =sample size and  $n$ =number of predictors. Since the Rosenberg model uses only one predictor, the adjusted  $R^2$  is not different from the original  $R^2$ . It can be seen that even for the Fishbein and the Sheth models the decrease in  $R^2$  due to adjustment is very minimal indicating that, especially in the case of the Sheth model, the correlation coefficients are not spuriously high due to the number of predictors.

### Discussion

The results clearly indicate that the Sheth model has a higher predictive validity than the Fishbein model. However, both models perform equally well in terms of cross-validation. The Rosenberg model was found to vary considerably in its predictive power over different samples. This points to the fact that although the Rosenberg and the Fishbein model are mathematically similar in the measurement of attitude, the nature of the constructs produces vastly different results. The better performance of the Fishbein model cannot be attributed to the inclusion of the normative beliefs component because it did not contribute significantly to the prediction of behavioral intention. The improved performance of the Sheth model over the Fishbein model, on the other hand, can be attributed to a more complete treatment of variables involved in the process, the factor analytic approach adopted and the measurement of behavioral intention incorporating needs and motives.

Apart from the relative predictive powers of the models, the study also provides other interesting conclusions.

- 1) Attitudes are effective predictors of behavioral intentions. This is confirmed by all the three models in this study. Further, both attitude toward the act (Fishbein model) and attitude toward the object (called predisposition in Sheth's model) prove to be significant indicating that either they are both important or that the distinction is not useful.
- 2) Anticipated situational influences are important for prediction of behavioral intention and behavior at least in the consumer behavior context. The fact that anticipated buying situation was a significant predictor of behavioral intention in the Sheth model suggests that buying an automobile might be largely influenced by the kind of deal that a person gets on a car or his expectation that some other alternative will become available before the time of purchase.
- 3) The poor predictive ability of social beliefs in both Fishbein and Sheth models is somewhat surprising in the situation of buying automobiles. Our conclusion is that Pinto is probably a universal car devoid of any stereotype and that the sample used was homogeneous with respect to life cycle, socioeconomic status and life style which might have minimized stereotype differences toward Pinto.

### Conclusion

This study has attempted to demonstrate the importance of comparing and cross-validating attitude models. The authors agree that one study alone, such as this, is not sufficient to prove conclusively the superiority of any model. The need for replication of such studies over a variety of behaviors and situations is, therefore, imminent. It is possible that one model might be appropriate in a buying situation whereas another might be appropriate in a social situation. Once the models are cross-validated, the next step would be to determine the extent to which they are applicable to different populations, which is labeled as 'validity generalization' in this study. Finally, it is hoped that the extensive validation and comparison of existing models will lead to a better understanding of consumer choice processes.

### FOOTNOTES

1. P.S. Raju and Rabi S. Bhagat are doctoral students in the Department of Business Administration and Jagdish N. Sheth is Illinois Business Associates Distinguished Professor of Business and Research Professor at the University of Illinois at Urbana-Champaign.
2. We thank Professor Martin Fishbein of the Department of Psychology, University of Illinois for his help in the wording and scale construction aspects of his model.

### REFERENCES

- Ajzen, I., & Fishbein, M. Attitudinal and normative variables as predictors of specific behaviors: A review of research generated by a theoretical model. Paper presented at the Attitude Research and Consumer Behavior Workshop, University of Illinois, Urbana, 1970.(a)
- Ajzen, I., & Fishbein, M. The prediction of behavior from attitudinal and normative variables. Journal of Experimental Social Psychology, 1970, 6, 466-87.
- Ajzen, I., & Fishbein, M. Attitudinal and normative variables as predictors of specific behavior. Journal of Personality and Social Psychology, 1973, 27, 41-57.
- Alpert, M.I. A canonical analysis of personality and the determinants of automobile choice. Combined Proceedings of the American Marketing Association, Spring and Fall Conferences, 1971, 312-16.
- Belk, R.W. An exploratory assessment of situational effects in buyer behavior. Journal of Marketing Research, 1974, 11, 156-63.
- Cohen, J.B., & Ahtola, O.T. An expectancy x value analysis of the relationship between consumer attitudes and behavior. Proceedings of the Second Annual Conference of the Association for Consumer Research, 1971, 344-64.
- Day, G.S. Evaluating models of attitude structure. Journal of Marketing Research, 1972, 9, 279-86.

- Fishbein, M. Attitude and the prediction of behavior. In M. Fishbein (Ed.), Readings in attitude theory and measurement. New York: John Wiley, 1967.
- Fishbein, M. The search for attitudinal behavioral consistency. In J. Cohen (Ed.), Behavioral Science foundations of consumer behavior. Glencoe, Ill.: Free Press, 1972.
- Howard, J.A., & Sheth, J.N. The theory of buyer behavior. New York: John Wiley, 1969.
- Kaplan, A. The conduct of inquiry: Methodology for behavioral science. San Francisco: Chandler Publishing Co., 1964.
- Katz, D. The functional approach to the study of attitudes. Public Opinion Quarterly, 1960, 24, 163-204.
- Katz, D., & Stotland, E. A preliminary statement to the theory of attitude structure and change. In S. Koch (Ed.), Psychology: A study of a science, Vol. 3. New York: McGraw-Hill, 1959.
- Lutz, R.J., & Howard, J.A. Toward a comprehensive view of the attitude-behavior relationship: The use of multiple set canonical analysis. Proceedings of the Social Statistics Section, American Statistical Association, 1971, 215-24.
- Miller, G.A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.
- Mosier, C.I. Problems and designs of cross-validation. Educational and Psychological Measurement, 1951, 11, 5-11.
- Rokeach, M. Beliefs, attitudes, and values: A theory of organization and change. San Francisco: Jossey-Bass, 1968.
- Rosenberg, M.J. Cognitive structure and attitudinal affect. Journal of Abnormal and Social Psychology, 1956, 53, 367-72.
- Rosenberg, M.J. A structural theory of attitude dynamics. Public Opinion Quarterly, 1960, 24, 319-40.
- Sandell, R.G. Effects of attitudinal and situational factors on reported choice behavior. Journal of Marketing Research, 1968, 5, 405-8.
- Sheth, J.N. An investigation of relationships among evaluative beliefs, affect, behavioral intention, and behavior. University of Illinois, College of Commerce and Business Administration. Working paper, 1970.
- Sheth, J.N. Affect, behavioral intention, and buying behavior as a function of evaluative beliefs. In P. Pellemans (Ed.), Insights in consumer and market behavior. Belgium: Numur University Publications Universitaires, 1971.

- Sheth, J.N. Reply to comments on the nature and uses of expectancy-value models in consumer attitude research. Journal of Marketing Research, 1972, 9, 462-65.
- Sheth, J.N. Brand profiles from beliefs and importances. Journal of Advertising Research, 1973, 13(1), 37-42.
- Sheth, J.N. A field study of attitude structure and attitude-behavior relationship. In J.N. Sheth (Ed.), Models of buyer behavior: Conceptual, quantitative, and empirical. New York: Harper and Row, 1974.
- Sheth, J.N., & Park, C.W. Equivalence of Fishbein and Rosenberg theories of attitudes. University of Illinois, College of Commerce and Business Administration. Working paper #108, 1973.
- Sheth, J.N., & Raju, P.S. Sequential and cyclical nature of information processing models in repetitive choice behaviors. Proceedings of the Fourth Annual Conference of the Association for Consumer Research, 1973, 348-58.
- Sheth, J.N., & Tuncalp, S. Structural assumptions underlying Fishbein's expectancy-value model of attitudes. University of Illinois, College of Commerce and Business Administration. Working Paper #207, 1974.
- Triandis, H.C. Attitude and attitude change. New York: John Wiley, 1971.
- Tuncalp, S. A study of the relationship between consumer's perceived quality judgements about a product based on the extrinsic and intrinsic product attributes and the underlying cognitive structure: An empirical investigation. Unpublished doctoral dissertation, University of Illinois, 1973.
- Wilkie, W.L., & Pessemier, E.A. Issues in marketing's use of multi-attribute attitudinal models. Journal of Marketing Research, 1973, 10, 428-41.



THE OBJECTIVE SITUATION AS A DETERMINANT OF CONSUMER BEHAVIOR<sup>1</sup>

Russell W. Belk<sup>2</sup>  
University of Illinois

Because the influence of situations is so pervasive, the notion that consumer behavior depends upon the situation is as easy to embrace as it is difficult to extricate. The development of taxonomies of consumer situations is discussed as one essential step in obtaining an understanding of the role of situations in determining consumer behavior. A means for developing such taxonomies based on similarities in behaviors elicited by objective situations is proposed and illustrated. Additional issues concerning the relationship between situations and behavior are outlined, and the relevance of objective characterizations of situations to the study of these issues is discussed.

The potential of situations to affect behavior presents a challenging premise for a new approach to the study of consumer behavior. There is now ample evidence and speculation that consumer behavior is a function of the interaction between the individual and the situation.<sup>3</sup> There is convincing reason to suspect that substantially more behavioral variance may be accounted for by introducing situational variables into research than by ignoring or controlling them.<sup>4</sup> And there is adequate justification for the conclusion that current explanations and predictions of consumer behavior offer much room for the improvement which may be provided by specifying situational conditions. What there is not, is any agreement as to just how the delineation and inclusion of situations in consumer research should proceed. It is the thesis of this paper that the objective characterization of consumer situations is a vital link in any such research process. After developing and illustrating this perspective, its place in an overall scheme for investigating situational influence on consumer behavior is elaborated.

#### The Concept of Situation

A situation is most simply defined by a locus in time and space. At this most basic conceptual level there is widespread, although not universal, consensus.<sup>5</sup> As Hansen (1972) points out, the situations of greatest concern in consumer behavior occur at times and places of communication, purchase, and consumption. Agreement also seems likely on this position. But beyond drawing these very general boundaries and categories for situations of interest, the conceptualization of situation is vague. W. I. Thomas began researching situation 50 years ago under a definition of situation as a configuration of factors conditioning a behavior reaction (Thomas, 1927). His work has since been criticised for giving rise to studies examining such wide ranging factors of "situation" as social institutions, groups, neighborhoods, individual experiences, events, and even complete illusions (Volkart, 1951). Other researchers have restricted their definitions of situation to include only role expectations and expected reactions of others (Warner and DeFleur, 1969), only geographical and architectural surroundings (Golledge and Zannaras, 1973), or only educational background (Nasiter, 1968). In attempting to operationalize Rokeach's concept



of situation, Roman (1969) has even used "situation" to mean brand of product. This represents an extreme flexibility in interpreting a term which Rokeach (1967) has used to mean stimuli as broadly conceived as events, activities, and social systems.

Rather than seek to resolve such differences with a highly specific and artificial distinction as to what constitutes a situation, it is more useful at this point to view situation as "all those factors particular to a time and place of observation which do not follow from a knowledge of personal (intra-individual) and stimulus (choice alternative) attributes, and which have a demonstrable and systematic effect on current behavior" (Belk, 1974). This definition specifies that the person, the situation, and the stimulus object are distinct sources of influence on behavior, and parallels such classifications as Helson's separation of organic, contextual, and focal cues in perception (Helson, 1964). Moreover the definition stipulates that of the nearly infinite number of cues surrounding a person at a given time and place, the situation of concern includes only those features which can be shown to influence behavior. Under this stipulation only those cues within the individual's immediate sphere of sensation need be considered as potential components of the situation. Within the framework of these minimum criteria for a situation, research may empirically answer the question of what elements constitute a consumer situation, according to the ability of these elements to alter consumer behavior.

#### Specifying Potential Elements of a Situation

In order to determine the situational conditions which do influence behavior, it is first necessary to specify the situational conditions which might influence behavior. To illustrate, suppose we randomly selected a person who is at this moment entering a department store in one of the major urban areas of the world. The task of specifying the situation is clarified to some degree by first deciding which potential determinants of behavior should be attributed to the person and stimulus objects. The person may be characterized by observable features (e.g., sex, approximate age, weight, and height) as well as unobservable features (e.g., personality, intellect, occupation, and skills). Such features may be regarded as situation-free person characteristics. For example, if the person were observed in a restaurant instead of a department store, such characteristics would not be affected. Of the many potential stimulus objects in the department store, suppose this person is now standing in front of a counter containing gloves. Because the gloves are now potential objects for behavior, they may be removed from the situation and treated as a separate source of influence. The characteristics of the gloves which may affect behavior are only those which this person is able to perceive directly (e.g., color, size, labeling, materials, and textures). Any added significance for a pair of gloves derives from an interaction of these characteristics with those of the individual or the situation.

All remaining potential behavioral influences in the above scenario are a part of the potential situation. We may externally note the most obvious elements of the situation in such features as temperature, time of day, persons present, transactions taking place, sounds, odors, and decor. Like stimulus object attributes, these physical qualities are variously perceived by the individual, but the situation itself may be described independently of these interpretations. There may however be other stimuli specific to this time and

place which reside within the individual and have no external correlates present. These unobservable features involve such momentary or episodic states of the individual as moods, plans, and purposes. For instance, it may well be important to know whether the department store shopper is taking a few moments while on a lunch break to select a gift for the birthday of a family member, or whether the person is in the store to get out of the rain, return a defective product, or shop for another general or specific type of product. Inasmuch as such momentary internal states may potentially affect behavior, they fall within the scope of the minimum criteria for a situation as long as they may not be construed as the result of an interaction between the characteristics of the individual and those of the stimulus object or the physical features of the situation. Thus a mood created by the background music in the store would be a response rather than a situational stimulus, while a mood created by events outside of the present physical situation may be regarded as a part of the current situational stimuli.

### The Objective Situation

Recognition of both external and internal aspects of the situation poses a major stumbling block to situational research. Mausner (1963) states the problem must succinctly: "If one specifies the stimulus in terms of the nature of the receiver, lawfulness becomes impossible. But without knowledge of the state of the receiver one cannot even begin to state the nature of the antecedent conditions which may be expected to lead to behavior" (Mausner, 1963, p. 107). For instance, the explanation that a person refrained from buying a new brand because the situation was a threatening one, is tautological if both the behavior and the situation are viewed as entirely ideosyncratic to the person. It tells us no more than the comparably hollow individual differences explanation that this is the sort of behavior engaged in by the sort of person who engages in this sort of behavior. Both explanations concentrate on the person and the response, and neglect the antecedent conditions for the behavior. While the inclusion of completely observable situational characteristics offers a chance for improved explainability, in this case such a description would fail to disclose the factors which make the situation threatening unless this condition would always result from the presence of the same set of physical stimuli. If instead the situation is threatening because of an internal mood or purpose of the individual, the richness added to the specification of situation by a knowledge of this state is lost.

There is a way in which we can maintain much of the potential richness of the concept of situation while still maintaining an essentially external perspective. To do so requires that the situation be operationally defined in "objective" terms. An objective element of the situation is one which is capable of external verification without the need to construct measures of internal states of the individual. In addition to the physical features of the situation, 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 about which friends or family have recently raised complaints, or 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 participant. Similarly we might ascertain that the individual has had a hectic day with the children,

or has just finished the last of several difficult final examinations, or has just gotten a promotion, instead of attempting to measure the mood this person brings to the situation. Besides the potential for increased explanatory richness compared to completely physically conceived situations, objective situational specifications allow greater accuracy in description and control of the situation than is possible using internal specifications.

#### Taxonomy of Consumer Behavior Situations

The discussion to this point has been concerned with a more precise specification of what is meant by the term "situation." It has been argued that a rich and meaningful use of this concept includes both the antecedent conditions for the momentary individual states which a person brings to a given time and place, and the physical features which he finds there. If we were able to fully enumerate the objective situations which this concept entails, we would doubtless conclude that no two situations are absolutely identical. Since the same may be said of consumers, the task of understanding the role of situations in determining behavior is no more or less imposing than that of understanding the role of consumer characteristics in determining behavior. The study of both of these sources of influence on consumer behavior may be aided substantially by the development of taxonomies which systematically classify their respective domains. However inadequate they may be, taxonomies of individual differences do exist in the form of psychological and behavioral classifications of individuals or attributes of individuals. There have been few attempts to develop comparable classifications of situations,<sup>6</sup> and no existing taxonomy is especially appropriate for describing consumer behavior situations. There is however a method of developing limited taxonomies of situations which is especially germane to the classification of consumer behavior situations. This method recognizes that rather than classifying situations on the basis of shared attributes, it is more directly relevant to classify situations on the basis of their ability to elicit similar behaviors. Such an approach simultaneously helps to reduce and group the many situations in which consumer behavior takes place, and does so in a way which fulfills the definitional criterion that a situation include only those elements which can be shown to systematically affect behavior.

#### A Method for Behavioral Taxonomy of Situations

The starting point for classifying situations via their tendency to elicit similar behaviors is a three dimensional data matrix where the dimensions are persons, situations, and behaviors. The data within this matrix are observations of whether or not a given behavior occurs for a given person in a given situation. Although this observational approach for collecting data for the matrix has been demonstrated to be feasible (Fredericksen, Jensen, and Beaton, 1972), a projective approach which asks subjects how likely they would be to engage in various behaviors in each of various situations has been more common. In either case, the three dimensional matrix which results may be reduced to a more general form in several ways. The traditional solution would be to collapse or average over one dimension to provide a two-way table on which to derive correlations for input to factor or cluster analysis. In order to classify situations in this manner the matrix would be averaged over subjects, and situations would be correlated across behaviors. In order to classify behaviors the matrix would be averaged over situations, and behaviors would be correlated across subjects. Factor analyzing these sets of correlations would

yield types of situations and types of behaviors, but it would not yield insights into the relationship between the two sets of factors obtained. That is, no description of types of behavior evoked by types of situations would result. In order to deal with this problem, Levin (1963) first demonstrated the application of Tucker's three-mode factor analysis to behavioral taxonomy (Tucker, 1963). Three-mode factor analysis of the sort of data suggested here not only provides classifications of types of behavior evoked by types of situations, it also distinguishes types of individuals for whom elicited behaviors show different patterns over the set of situations. The procedure simultaneously extracts factors in each of the three modes of the matrix, and describes their interrelationships in terms of a core matrix. By multiplying the entries in this three dimensional core matrix by the entries in the three factor loading matrices before rotation or factor deletion, the original data matrix may be reconstructed. Summary details of this technique may be found in Vavra (1972a,b) and Belk (1973,1974).

#### Taxonomy of Objective Consumer Situations -- An Example

In order to apply three-mode factor analysis to the development of situational taxonomies, an illustrative set of consumer situations and choice behaviors were prepared concerning patronage of fast food and take-out restaurants.<sup>7</sup> Ten situations and ten alternative choices were presented to subjects who responded on a six point scale from "Not at All Likely" to "Very Likely" to indicate the probability that each choice would be made in each situation. Subjects were 98 married females in a single community. Their responses to each stimulus and response pair were standardized, and the transformed data were analyzed using three-mode factor analysis.

The situations employed and their loadings on the four factors derived in this mode are shown in Table 1. The selection of situations for investigation cannot hope to be representative until further applications of behavioral taxonomies have been accumulated. The situations used in this application may be seen to be specified in objective terms, with more emphasis placed on the states the individual brings to the situation than on the physical features he finds there. Based on the factor loadings in Table 1, it appears that the first situation factor may be described as variety-seeking situations. The second factor involves situations in which the subject is entertaining guests. Factor three is clearly picnic situations. And the final factor seems to represent occasions which are relaxation-seeking and informal. The last factor tends to include situations which occur at the end of a long day or upon completion of some involving activity. Interestingly, whether these antecedent conditions tend to be positive (a pleasant chat; having too good a time) or negative (too tired) their effect on choice behavior is similar.

The other modes analyzed provided four factors for responses and four factors for persons. The response factors described 1) multi-course dinners, 2) fast food A, 3) fast food B, and 4) self-prepared hot dogs or sandwiches. The relationship of these choices to the four types of situations is shown in Table 2. This table is the portion of the core matrix corresponding to one of the four types of persons, and the entries reflect the relative likelihood of choosing each alternative meal type under each alternative situation type for this subgroup of people. The more negative entries indicate unlikely choices and the more positive entries reflect likely choices. The type of person depicted is most likely to select takeout food A in variety-seeking situations

TABLE 1  
 Factor Loadings of Situations  
 (From Three-mode Factor Analysis with Varimax Rotations)

Situation	I	II	III	IV
1. You are too tired to cook dinner either because you have been cleaning the house all day or you had a very busy day at the office or you had been shopping all day or the children have given you a hectic day.	.07	.02	-.02	.47
2. Your neighbor came over to visit and you are having a pleasant chat, and you discover it is lunch time.	-.67	.23	-.03	.35
3. You are having a few friends over for a causal get-together.	-.11	.69	-.03	-.01
4. You are planning a picnic with just your own family.	-.02	-.02	.69	.04
5. You have been watching the afternoon movie or you have been playing with the children, and they don't want you to stop the game, or you have been playing tennis or cards and you are having too good a time to cook dinner.	.02	.04	.05	.47
6. You want a change from the daily routine.	.45	.26	-.09	.15
7. You are wondering what to serve yourself, and the children since your husband is not going to be home for dinner.	.01	-.19	.02	.58
8. You are planning a picnic with your friends.	.01	.03	.71	-.02
9. You need food for some unexpected dinner guests.	.16	.59	.07	-.07
10. The children are asking for something different for dinner.	.54	.03	.03	.23

and relaxation-seeking situations, and takeout food B for picnics and when entertaining guests. Others of the four person types derived showed markedly different patterns of behaviors. The multi-course dinners tended to be unlikely choices for all groups, which is consistent with the observation that these were generally less popular fast food outlets in the community. Twenty of the 98 subjects were classified as belonging to the classification of persons shown, and they were further characterized by their higher age, fewer children, and lower income than those in the other three groups.

This example serves to suggest the potential uses of behavioral taxonomies for situations involving a particular product category. Obviously it would not be suggested that the taxonomy derived here will necessarily describe relevant situations in other product or behavior categories. Of the four types of situations found to affect choices in this instance, only the relaxation-seeking factor finds parallels in similar studies for choice of snack products, meat products, and motion pictures (Belk, 1973,1974). The

TABLE 2  
CORE MATRIX FOR PERSON FACTOR ONE  
(From Three-mode Factor Analysis with Varimax Rotations)

Response Factors	Situation Factors			
	Variety- Seeking	Having Guests	Picnics	Relaxation- Seeking
Full Course Dinner	-1.31	-10.07	-11.36	-13.12
Takeout Food A	5.93	5.14	-5.58	9.66
Takeout Food B	4.25	10.47	9.69	1.58
Self-prepared	-12.20	-5.41	8.93	5.49

technique of behavioral taxonomy, however, is highly generalizable. It presents a useful and versatile means of screening potential stimulus configurations for their behavioral impact, while at the same time beginning the process of categorizing and compressing these situations. The full results of analyses like the one illustrated provide nearly complete formulas (types of choices by types of consumers in types of situations) for situational segmentation and positioning. Where the researcher has been careful to specify input situations in objective terms, the further necessary information on frequency of situational occurrences for each type of individual may be readily obtained as well.

#### The Role of Objective Characterizations in Situational Research

Although obtaining behavioral taxonomies of objectively specified consumer situations is a concrete means of operationalizing the concept of situation, such studies are only the starting point for investigating situational influence. Once it has been demonstrated that a set of situations have the capability to influence consumer behavior in a certain manner, the next obvious question is why this effect occurs. In light of the core matrix from three-mode factor analysis, we may also ask why this effect occurs for some persons but not for others. To answer both questions, and especially the latter, it will be necessary to consider the internal responses which an individual makes to a situation. Examinations of the links between objective situations and internal responses are needed to reveal the ways in which the individual's thoughts, feelings, and expectations are altered by particular situations and situational cues. These internal responses, to the extent they can be measured, may then be scrutinized to form causal hypotheses about the nature of situational effects. This type of research may be expected to suggest explanations for effects which are revealed by direct situation-behavior research but nevertheless appear mysterious. For instance, the fact that both pleasant and unpleasant antecedent activity lead to similar behaviors in the prior example, may find some rationale in terms of similar expectations generated by such activities. In carrying out such research, the link between the situation as it objectively exists and the situation as it is subjectively interpreted forms an anchor to the real world which is essential if we are to avoid the sort of tautologies mentioned earlier. It is most likely that it is the absence of this crucial link in Lewin's (1935) separation of life space from physical space which has

impeded the literal adoption of this perspective, despite its substantial conceptual contributions. The warning then is to begin with the situation as it exists rather than its internal representation.

Another area of concern for consumer situational research is how an individual gets to be a particular situation in the first place. Again a knowledge of internal responses to situations may help. The degree to which a given situation is attractive, avoidable, and foreseeable, should be related to its frequency of occurrence, which should in turn determine the aggregate frequency of the behaviors evoked by this situation. Research on both perception of situations (Mehrabian and Russell, 1974) and the frequency of situational occurrences (Szalai, 1966) is now under way, although the combination of these two research streams in terms of the individual's control in selecting the situations he encounters has not yet taken place. In order for a merger of these two perspectives to be possible, they must have a common denominator. Only if the situations toward which perceptions are measured are the same situations for which frequency distributions are constructed, can the two be related. In order to tie both research streams to the same objective world, the development of taxonomies of objective situations is a high priority task.

One further and more obvious contribution of objective characterizations of consumer situations involves the implications flowing from observations of situational influence on behavior. From the points of view of marketing and public policy, the relevant question is "how externally controllable are those elements of consumer situations which can be shown to influence behavior." It is apparent that the question of control is more easily assessed where the situation is measured in terms of objective elements than where the focus is on subjective interpretations. For instance a retail shopping situation described by such factors as merchandise clutter and floor surfacing has more readily apparent implications than a description of the situation by its novelty and arousal potential. Furthermore internal responses such as mood, that are only partially controllable, can be more adequately evaluated for controllable (e.g., color) and uncontrollable (e.g., weather) factors when they are specified in objective terms.

#### Conclusion

The clearest means of defining consumer situations resides in the objective features which characterize a locus in time and space. Such elements include both antecedant conditions for the internal states the participant brings to the situation and the physical elements he attends there. These features lend themselves to direct behavior taxonomy without the need to infer internal responses of the individual to the situation. While objectively defined situations are only one component of the necessary perspectives for researching situations, the objective situation is the common thread which is required of any study of situations if it is to add to our ability to explain, predict, or alter consumer behavior.

## FOOTNOTES

1. Portions of this research utilize data supplied by Leo Burnett, U.S.A.. The cooperation of Stuart Agres in making these data available is gratefully acknowledged.
2. Russell Belk is an assistant professor in the College of Commerce and Business Administration, University of Illinois, Urbana, Illinois 61801.
3. Speculation on this point is clearest by Ward and Robertson (1973), Engel, Kollat, and Blackwell (1969), and Tucker (1964). Several seminal studies have found the influence of situational interactions to be more sizeable than the variance attributable to the influence of individual differences alone (Sandell, 1968; Bishop and Witt, 1970; Belk, 1974).
4. This hypothesis is supported by the studies cited in footnote 3 and in studies by Hansen (1972), Sheth (1971), and Endler and Hunt (1969).
5. A primary exception is Barker's (1968) "behavioral setting" which adds a behavioral dimension to this locus. By using a behavioral sequence or "action pattern" to help in defining boundaries for the situation, Barker extends the time and place dimensions to potentially longer and larger units than those considered in the following discussion. By extending all three dimensions to still larger units, we approach the concept of environment, as it is coming to be used (see Proshansky, Ittelson, and Rivlin, 1970).
6. The several notable attempts to develop comprehensive taxonomies of situations are those by Sherif and Sherif (1956), Sells (1963,1969), Bellows (1963), Allen (1965), and Kasmar (1970).
7. Data were provided by Stuart J. Agres, Manager of Research Development, Leo Burnett.

## REFERENCES

- Allen, B. L. Situational factors in conformity. In L. Berkowitz (Ed.), Advances in experimental social psychology. Vol. 2. New York: Academic Press, 1965.
- Barker, R. G. Ecological psychology: Concepts and methods for studying the environment of human behavior. Stanford: Stanford University Press, 1968.
- Belk, R. W. Application and analysis of the behavioral differential inventory for assessing situational effects in consumer behavior. In S. Ward and P. Wright (Eds.), Proceedings. Urbana, Ill.: Association for Consumer Research, 1973, pp. 370-380.
- Belk, R. W. An exploratory assessment of situational effects in buyer behavior. Journal of Marketing Research, 1974, 11, 156-163.
- Bellows, R. Toward a taxonomy of social situations. In S. B. Sells (Ed.), Stimulus determinants of behavior. New York: Ronald Press, 1963.



- Bishop, D. W. and Witt, P. A. Sources of behavioral variance during leisure time. Journal of Personality and Social Psychology, 1970, 16, 352-360.
- Endler, N. S. and Hunt, J. McV. S-R inventories of hostility and comparisons of the proportions of variance from persons, responses, and situations for hostility and anxiousness. Journal of Personality and Social Psychology, 1969, 9, 309-315.
- Engel, J. F., Kollat, D. T., and Blackwell, R. D. Personality measures and market segmentation. Business Horizons, 1969, 12, 61-70.
- Fredericksen, N., Jensen, O., and Beaton, A. Prediction of organizational behavior. Elmsford, N.Y.: Pergamon Press, 1972.
- Golledge, R. G. and Zannaras, G. Cognitive approaches to analysis of human spatial behavior. In W. H. Ittelson (Ed.), Environment and cognition. New York: Seminar Press, 1973.
- Hansen, F. Consumer choice behavior. New York: The Free Press, 1972.
- Helson, H. Adaptation-level theory. New York: Harper and Row, 1964.
- Kasmar, J. V. The development of a usable lexicon of environmental descriptors. Environment and Behavior, 1970, 2, 153-169.
- Levin, J. Three-mode factor analysis. (Doctoral dissertation, University of Illinois) Ann Arbor, Mich.: University Microfilms, 1963. No. 6099.
- Lewin, K. A dynamic theory of personality. New York: McGraw-Hill, 1936.
- Mausner, B. M. The specification of the stimulus situation in a social interaction. In S. B. Sells (Ed.), Stimulus determinants of behavior. New York: Ronald Press, 1963.
- Mehrabian, A. and Russell, J.A. An approach to environmental psychology. Cambridge: The MIT Press, 1974.
- Nasiter, D. A note on the contextual effects and the political orientation of university students. American Sociological Review, 23, 1968, 210-214.
- Proshansky, H. M., Ittelson, W. H., and Rivlin, L. G. Environmental psychology. New York: Holt, Rinehart, and Winston, 1970.
- Rokeach, M. Attitude change and behavior change. Public Opinion Quarterly, 30, 1967, 529-550.
- Roman, H. S. Semantic generalization in formation of consumer attitudes. Journal of Marketing Research, 1969, 6, 369-373.
- Sandell, R. G. Effects of attitudinal and situational factors on reported choice behavior. Journal of Marketing Research, 1968, 4, 405-408.
- Sells, S. B. Dimensions of stimulus situations which account for behavioral variance. In S. B. Sells (Ed.), Stimulus determinants of behavior. New York: Ronald Press, 1963.

- Sells, S. B. Toward a taxonomy of organizations. In W. W. Cooper, A. J. Leavitt, and M. W. Shelly II (Eds.), New perspectives in organization research. New York: Wiley, 1969.
- Sheth, J. N. Canonical analysis of attitude - behavior relationship. Paper presented at TIMS eighteenth international meeting, 1971.
- Sherif, M. and Sherif, C. W. An outline of social psychology. Rev. Ed. New York: Harper and Row, 1956.
- Szalai, A. The multinational comparative time budget research project. American Behavioral Scientist, 1966, 10, 1-12+.
- Thomas, W. I. The behavioral pattern and the situation. Proceedings, Twenty-second annual meeting, American Sociological Society, 22, 1927, 1-13.
- Tucker, L. R. Implications of factor analysis of three-way matrices for the measurement of change. In C. W. Harris (Ed.), Problems in measuring change. Madison: University of Wisconsin Press, 1963.
- Tucker, W. T. The social context of economic behavior. New York: Holt, Rinehart, and Winston, 1964.
- Vavra, T. G. An application of three-mode factor analysis to product perception. Proceedings, American Marketing Association fall conference, 1972a, Chicago: American Marketing Association, 578-583.
- Vavra, T. G. Factor analysis of perceptual change. Journal of Marketing Research, 1972b, 9, 193-199.
- Volkart, E. H. Introduction: Social behavior and the defined situation. In E. H. Volkart (Ed.), Social behavior and personality. New York: Social Sciences Research Council, 1951.
- Ward, S. and Robertson, T. S. Consumer behavior research: Promise and prospects. In S. Ward and T. S. Robertson (Eds.), Consumer behavior: Theoretical sources. Englewood Cliffs, N.J.: Prentice-Hall, 1973.



THE PSYCHOLOGICAL SITUATION AS  
A DETERMINANT OF CONSUMER BEHAVIOR<sup>1</sup>

Richard J. Lutz and Pradeep Kakkar<sup>2</sup>  
University of California, Los Angeles

The effects on behavior of the situation in which that behavior occurs have long been recognized, but seldom systematically investigated. The present study reviews some of the recent research on situational effects in marketing and suggests possible extensions. Specifically, a theory developed in environmental psychology by Mehrabian and Russell is applied in a partial replication of Belk's recent research in consumer psychology. This new framework relies on three internal state variables--pleasure, arousal, and dominance--which are conceptualized as mediating the influence of the situation on behavior. Present results indicate that the Mehrabian-Russell framework may be quite useful in the understanding and classification of consumption-related situations.

The effects of "the situation" on behavior have been alluded to frequently in the psychological and marketing literatures. Many articles include situational qualifiers in their Discussion sections, attempting to delineate those situations over which the observed relationships may be generalizable. Yet, as noted by Cottrell (1950), the actual usage of the term "situation" has been "... quite as great in preserving an illusion of understanding as it is in conveying genuine comprehension" (p. 711). Despite this rather skeptical assessment of the status of situational variables, Cottrell went on to advocate the necessity for a situational perspective in the study of human motivation. Similarly, Frank, Massy and Wind (1972) have recently pointed to the value of "situation-specific" variables in market segmentation. Thus, it appears that more systematic investigation of situational variables would make a valuable contribution to both psychology and marketing.

There have been a few scattered articles on situational influence in the marketing literature (e.g., Sandell, 1968; Gronhaug, 1972; Hansen, 1972), but only recently has any systematic empirical research been undertaken (Belk, 1974a,b,c). The purpose of the present paper, therefore, is primarily developmental; drawing selectively from recent literature pertaining to situational influence on behavior, recommendations will be made regarding potential directions for situational research in consumer psychology.

#### Defining the Situation

A crucial condition which must be met in order to allow progress in situational research is an adequate conceptual definition of the phenomenon. One key distinction has been offered by Belk (1974c):

....situation and environment ...represent distinct sources of influence on consumer behavior and should not be used synonymously. Environment is the broader construct and represents a general milieu of behavior, whereas situation is a more momentary concept (pp. 1-2).

This distinction appears meaningful, useful, and operationalizable and should be maintained in future research. Having made this distinction, Belk (1974b) further defined "situation:"

This study adopts a general view of situation as something outside the basic tendencies and characteristics of the individual, but beyond the characteristics of the stimulus object to be acted upon.....situation may then be defined as all those factors particular to a time and place of observation which do not follow from a knowledge of personal (intra-individual) and stimulus (choice alternative) attributes, and which have a demonstrable and systematic effect on current behavior. It should be noted that this definition describes situation in terms of observable aggregate effects rather than in terms of similarities in individual perceptions of situations (pp. 156-7).

In restricting the definition of situation to "observable aggregate effects," Belk differs with most situational theorists. For instance, Cottrell (1950) stated (rather categorically): "We are quite certain that individuals and groups react to their own definitions of situations..." (p. 711), thus suggesting a subjective, rather than objective, definition. Similarly, Rotter (1955) states: "The basic principle for classifying or categorizing a situation is psychological, that is, subjective..." (p. 259), and relates it to Lewin's concept of the "life space." Murray (1952), in his theory of situational press, suggested classification of situations based on the differing effects they exert on the individual. Finally, Hansen (1972) concludes: "Altogether, how the actor perceives the situation is as important as the actual elements found in the physical environment" (p. 47).

On the other hand, Barker (1968) and his associates have developed a theory of "behavior settings" which allows specification of situations based upon only physical characteristics and observable behavior patterns, with no consideration of internal factors. The cataloging of an almost infinite array of physical stimuli has been a major drawback to this approach, however, and has been likened to the similarly impossible task of deriving a complete list of human motives. Even among those researchers who have advocated an objective approach to the study of situational influence, a strong emphasis has been placed upon subjective factors as well. Volkart (1951) stated: "The human situation often includes some factors common to both the observer and the actor...[but] also includes some factors that exist only for the actors, i.e., how they perceive the situation, what it means to them, what their 'definition of the situation' is" (p. 2). Similarly, Thomas (1937) posited that: "An adjustive effort of any kind is preceded by a decision to act or not act along a given line, and the decision is itself preceded by a definition of the situation, that is to say, an interpretation, or point of view...." (p. 8), and that to adequately describe a situation, the situation must be studied "...as it exists in verifiable, objective terms, and as it has seemed to exist in terms of the interested persons" (cited in Volkart, 1951, p. 6).

Thus, there appears to be a definite need to adopt a subjective view of the situation in order to understand its effects on human behavior. As Hansen (1972) succinctly stated:

Even if it were found that situational variations explain all variations in behavior, the question must still be asked how the individual transforms the situational input to behavioral output, which makes it necessary to look at the individual's

internal processes (p. 42).

Therefore, the present research adopts the viewpoint that the "situation" relevant for the understanding of consumer behavior is the psychological situation, which may be defined as an individual's internal responses to, or interpretations 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.

The above definition is similar in many respects to the one offered by Belk (1974b), but differs in the focus on subjective interpretation of the situation. Additionally, psychological processes preceding behavior are posited as a locus of situational effects, rather than overt behavioral outcomes only. This process orientation has been adopted previously by Wright (1973, 1974) and would seem to be an important area of enquiry, given the current interest in consumer information processing. In essence, the present approach is not intended to be a competitor, but rather a complement, to Belk's earlier work, in which he acknowledged that various components of the situation "...gain meaning and effect only through the perceptions of the individual" (1974c, p. 5). The present definition simply modifies his earlier one to include explicit consideration of perceptual factors and the idea of process variables.

#### Need for a Taxonomy of Situations

Most situational theorists have recognized the need for a taxonomy of situations or situational components which would allow generalizable investigations of situational influence. Rotter (1955) suggested that "systematic schemata" be used to predict human behavior from knowledge of the situation and Moos (1973, 1974) and his associates have determined the dimensionality of eight different "environments" in attempting to develop a classification scheme. Other researchers have variously specified time (Cottrell, 1950; Wright, 1973; Belk, 1974c), group opinion (Gorden, 1952), physical surroundings (Barker, 1960; Belk, 1974c; Toffler, 1970), and goal structure (Gronhaug, 1972; Belk, 1974c) as aspects of situations. However, as noted by Volkart (1951), "...situations are multidimensional, and it is difficult to specify all their elements" (p. 19).

Yet it is clear that some sort of taxonomy is necessary in order for situational variables to become meaningful in the explanation of consumer behavior. A recent theory developed in the discipline of environmental psychology may represent the first tentative step in the pursuit of such a taxonomy. Environmental psychology is normally concerned with issues such as the effects of noise, color, temperature and other specific physical stimuli on behavior. Within this tradition, however, Mehrabian and Russell's (1974) proposals relating to the description of environments and the measurement of environmental impact provide a rich framework for the study of the general psychological situation.

#### The Mehrabian-Russell Theory

The fundamental proposition of Mehrabian and Russell's theory is that the impact of the situation on behavior is mediated by emotional responses, so that any set of conditions initially generates an emotional (affective, connotative, feeling) reaction, which in turn leads to a behavioral response. Further, the universe of all possible emotional responses may be represented by one or a combination of three basic dimensions: pleasure, arousal and dominance. Pleasure as an emotional state is distinguished from "preference, liking, positive

reinforcement, or approach-avoidance...since the latter responses are also determined by the arousing quality of a stimulus" (Mehrabian and Russell, 1974, p. 18). It is a composite of feelings such as happiness, contentment, satisfaction, etc. Arousal is an activity orientation and is "a measure of how wide awake the organism is, of how ready it is to act" (Berlyne, 1960, p. 48). Finally, dominance is a reflection of the extent to which the individual feels in control of or overpowered by his environment. The higher the level of dominance perceived in the situation, the more submissive is the state of the individual.

Among the extensive evidence offered by Mehrabian and Russell to substantiate their fundamental proposition is the body of literature relating to the semantic differential technique. There is a close parallel between the dimensions discussed by Mehrabian and Russell and by Osgood, Suci and Tannenbaum (1957): pleasure corresponds roughly to evaluation, arousal corresponds to activity, and dominance relates (inversely) to potency. The crucial element common to both frameworks is that the dimensions are not specific to modalities; i.e., the response dimensions are not anchored to any particular channel of experience, thereby allowing comparison of responses to varied types of stimuli.

To develop measures for the three mediating emotional variables, Mehrabian and Russell presented a series of several hundred hypothetical situations to subjects in three separate studies and obtained responses on several semantic differential scales designed to measure the three basic emotions described earlier. By factor analyzing the data, it was found that three factors corresponding to pleasure, arousal and dominance could be identified. The six items loading highest on each dimension were used to construct a simple 18-item instrument (shown in Appendix A) for self-report measurement of emotional response to any conceivable situation. For instance, a boring situation should elicit responses low in pleasure, arousal and dominance; a hostile situation, on the other hand, may generate a high score on arousal, a low score on pleasure, and a low score on dominance.

According to Mehrabian and Russell, not only can emotional response to any situation be represented by these three factors, but these same factors can also be used to categorize the situations which generate those emotional states. In other words, stimuli can be described according to the responses they generate. Thus an individual shopping hurriedly for a party to which he has invited some important guests might be described as being in a high arousal situation, in contrast to a virtually endless description of the observable situation. All physical and psychological stimuli are seen as resulting in some combination of the three basic emotions, so that the consumer's ultimate overt response is, at least in part, determined by these intermediate responses. Defining situational stimuli in terms of responses is totally consistent with the definition of the psychological situation presented earlier in this paper.

Not only does the Mehrabian-Russell approach allow the description of individual situations, it also facilitates the vital task of comparing different situations. Since any psychological situation can be measured along these three dimensions, situations can be compared through an analysis of the corresponding coefficients (or scores). This can be done at the individual level or for aggregates of individuals, so that conceivably one could obtain reactions of various market segments to consumption-related situations and thereby draw implications for marketing strategy.

Application of the Mehrabian-Russell theory to the study of the psychological

situation in consumer psychology, then, offers potential solutions for two of the pressing problems of situational research. On the one hand, the Mehrabian-Russell framework can be viewed as at least a partial explanation of the influence of situations on behavior, whereas most past research has utilized a "black-box" approach, focusing only on inputs and outputs, with no consideration of mediating psychological variables. Secondly, the Mehrabian-Russell framework may satisfy the need for a taxonomical scheme for situations. Thus, the remainder of this paper will be devoted to an empirical application of this framework in an attempt to assess its potential usefulness in the explanation and classification of situational influence.

## Method

### Partial Replication

In order to facilitate the evaluation of the Mehrabian-Russell framework, the present research was designed as a partial replication of Belk's (1974b) study of snack situations. The intent was to show how application of this new theory could lead to richer insights into the phenomenon of situational influence.

Belk (1974b) presented 100 subjects with ten different snack consumption situations, asking them to respond to a 10-item behavioral differential inventory in each situation. To determine the effects of situations, products, and individual differences, he then conducted a 100 (persons) x 10 (products) x 10 (situations) mixed effects ANOVA. The present research utilized the same situation descriptions and the same behavioral differential, but did not present every situation to each subject, thus ruling out the comparison of individual differences.

### Situation Descriptions

Both Belk (1974b) and Mehrabian and Russell (1974) used short paragraphs to describe situations to subjects. For the purposes of comparison, ten situations from Mehrabian and Russell were included in this research, in addition to Belk's ten snack situations. Paragraphs defining the twenty situations are presented in Tables 3 and 4.

### Measurement

Following each situation description, subjects responded to the 18-item PAD instrument. For the ten snack situations, the behavioral differential instrument described by Belk was also included. For the ten "general" situations taken from Mehrabian and Russell, no behavioral differential was included, since the items would have made no sense in those situations (Table 2 shows the ten behavioral differential items.).

### Procedure

Subjects were 315 undergraduate students at UCLA and California State University at Northridge. Each subject responded to one "general" situation and one "snack" situation. Treatment booklets were constructed so that the order and frequency of presentation of all twenty situations were balanced. The booklets were distributed, completed, and collected during a regular class session.



## Analysis and Results

Comparison with Belk's Findings

In order to provide at least a partial comparison of the current results with Belk's (1974b) earlier findings, a 10x10 (Situations x Products) fixed effects ANOVA was conducted on subjects' behavioral differential ratings.<sup>3</sup> Each S's responses were in the form of stated intentions to purchase each of 10 snack products in one of the 10 purchase situations. Thus, a given subject's responses corresponded to one level of the Situations factor and all ten levels of the Products factor. Results of the ANOVA are displayed in Table 1, together with the corresponding results from Belk (1974b).

TABLE 1

Relative Contributions to Variance in Behavioral Intentions

Source	df	Mean Square	F	Percent Contribution <sup>a</sup>	Percent Contribution <sup>b</sup>
Situations (S)	9	10.69	5.77 <sup>c</sup>	1.18	1.12
Products (P)	9	50.16	27.09 <sup>c</sup>	6.45	8.65
S x P	81	6.65	3.59 <sup>c</sup>	5.76	15.82
Residual	3050	1.85		86.61	44.37

<sup>a</sup>Calculated through the use of Hays' (1964) omega-squared statistic

<sup>b</sup>From Belk (1974b, Table 2, p. 159).

<sup>c</sup>Significant beyond  $p < .001$

As shown in the Table, results for the two main effects were quite similar in magnitude in the two studies. Situations explained very little variance, while products explained somewhat more across situations. While Belk does not present cell mean values, the means shown in Tables 2 and 3 for Behavioral Intention suggest that only two or three treatment levels on each factor are accounting for most of the two main effects.

The interaction of Situations and Products accounted for considerably less variance than reported by Belk. This may in part be due to the differences in methods employed in the two studies, as subjects did not respond to each of the ten situations in the present research. Other research reported by Belk (1974c) has tended to show a rather large Situation x Product interaction for beverages, meats and leisure activities; therefore, the present results may be reflecting some method variance. On the surface, there is no apparent reason why the present method should necessarily lead to a suppression of the interaction effect. In fact, the most plausible interpretation of the discrepancy in results may be that demand characteristics were operating in Belk's (1974b) experiment, causing subjects to exaggerate supposed shifts in consumption behavior across situations.<sup>4</sup> Careful attention should be focused on this point in future development of methods for studying situational influence.

Use of the PAD Instrument

In order to assess the stability of the PAD instrument, scores for the ten

TABLE 2

Mean Behavioral Intention Ratings for Ten Snack Products<sup>a</sup>

Product	Mean	Product	Mean
Fresh Fruit	3.79	Assorted Nuts	3.20
Cheese	3.54	Ice Cream	3.09
Potato Chips	3.38	Pastries	3.04
Sandwiches	3.34	Crackers	2.86
Cookies	3.25	Popcorn	2.32

<sup>a</sup>On 5-point scale from "not at all likely" (1) to "extremely likely" (5).

"general" situations in the present study were compared with scores for the same ten situations as reported by Mehrabian and Russell (1974). Table 4 shows the ten situation descriptions, together with group mean scores on Pleasure, Arousal and Dominance. Numbers in parentheses are the corresponding values reported in Mehrabian and Russell. Correlation coefficients computed between the present scores and the earlier ones revealed that the instrument is fairly stable, particularly for the Arousal measure ( $r=.99$ ). Pleasure ( $r=.86$ ) was somewhat less stable, with Dominance ( $r=.82$ ) being the least stable of the three dimensions.

Intercorrelations among the three dimensions computed across the entire sample ( $n=315$ ) showed that, while all three correlations were statistically significant, the strength of association was quite low ( $r_{PD}=.16$ ;  $r_{AD}=.20$ ;  $r_{AD}=-.13$ ). Thus it was concluded that the PAD instrument was performing in accordance with expectations, and further data analysis was undertaken.

Figure 1 displays the positions of the twenty situations investigated on the Pleasure, Arousal and Dominance dimensions, in the form of data cubes. Figure 1A shows the stability of scores for the "general" situations, while Figure 1B offers a comparison of the ten general situations with the ten "snack" situations. More polarization of response is evidenced on all three dimensions for the general situations, as compared with the snack situations. While this finding can hardly be regarded as evidence for a general phenomenon, it does serve to illustrate the possibility that purchase situations may be low involvement situations to the extent that they do not generate strong emotional reactions. This would support other research which has shown products to be relatively unimportant as compared with issues (Hupfer and Gardner, 1971). Nevertheless, the current situations were not randomly sampled, and as conclusion can be drawn regarding the issue of involvement, or more accurately, response polarity.

#### Relating Situational and Behavioral Responses

One of the primary advantages of the Mehrabian-Russell framework is that it allows quantification of situational descriptions along the three emotional response dimensions, thus facilitating the use of powerful multivariate statistical procedures in assessing situational influence. That is, rather than treating the ten snack situations in a nominal fashion (as in the ANOVA procedure), interval scale values can be attached to certain situational dimensions (i.e., pleasure, arousal, and dominance).

To measure the degree to which the three proposed mediating variables (P, A, and D) were related to stated behavioral intentions, a canonical correlation was

TABLE 3

Mean Behavioral Intention Ratings and PAD Instrument  
Scores for Ten Snack Situations

Situation Description	n	Intent	Pleasure <sup>a</sup>	Arousal <sup>a</sup>	Dominance <sup>a</sup>
11. You suddenly realize that you have invited a couple of friends over for the evening and you have nothing for them to snack on.	33	3.28	1.72	1.11	0.01
12. You are planning a party for a few close friends and are wondering what to have around to snack on.	33	3.54	0.13	1.39	-0.03
13. You are shopping for a snack that you or your family can eat while watching television in the evenings.	27	3.06	1.24	0.35	0.22
14. You are at the grocery store when you get an urge for a between-meal snack.	31	2.85	0.17	1.33	-0.07
15. Snacks at your house have become a little dull lately and you are wondering what you might pick up that would be better.	36	3.08	1.18	0.22	0.42
16. You are going on a long automobile trip and are thinking that you should bring along some snack to eat on the way.	33	3.19	1.15	0.59	-0.15
17. You are at the supermarket and notice the many available snack products; you wonder if you should pick something up in case friends drop by.	29	3.21	-0.03	0.20	0.02
18. You are thinking about what type of snack to buy to keep around the house this weekend.	32	3.09	0.72	0.05	0.16
19. You are at the store to pick up some things for a picnic you are planning with friends and are trying to decide what kind of snack to buy.	32	3.17	1.15	0.03	0.16
20. You are thinking about a snack to have with lunch at noon.	29	3.34	1.47	1.05	0.29

<sup>a</sup>Possible range of values from -4.00 to +4.00

TABLE 4

Mean Scores on Pleasure, Arousal and Dominance for  
Ten "General" Situations<sup>a</sup>

Situation Description <sup>b</sup>	n	Pleasure	Arousal	Dominance
1-A. You have just plunged into a cool pool at a public part on an extremely hot summer day. The pool is crowded. Many of the people around you are playing with a large rubber beach ball. The ball hits your head and bounces off a couple of times.	32	0.13 (0.02)	1.13 (1.10)	-0.09 (0.37)
2-B. You have been on a trip and are sitting along in your camper late at night in a roadside rest stop.	34	1.13 (0.58)	-0.49 (-1.20)	0.24 (0.18)
3-C. Imagine sitting by yourself in your own room, when it is quiet.	32	1.78 (1.16)	-0.67 (-1.05)	0.51 (0.90)
4-D. You are sitting on a small wooden bench in a cold, deserted classroom. You had attended first grade here. You have been sitting here for a few hours.	30	0.56 (0.44)	-0.38 (-0.23)	-0.36 (0.03)
5-E. You are teaching in a children's day care center. All the boys are running, and hitting each other, and screaming; a little girl sits a few feet away from them, banging on a drum.	32	-0.22 (-0.75)	1.76 (1.52)	0.08 (0.84)
6-F. You are with a scientific expedition in the Amazon jungle. Thick vines climb into the tall trees. The lush vegetation is so tall and thick you must fight for every step. Occasionally you glimpse the towering mountains in the distance. It is warm, and the wild birds chirp merrily.	30	0.91 (0.46)	1.69 (1.54)	-0.53 (-0.22)
7-G. Imagine yourself sitting alone in your own back yard on a quiet day when there isn't much of anything going on.	31	1.76 (0.57)	-1.00 (-1.84)	-0.14 (0.13)
8-H. You are taking a ride on a roller coaster. First it moves up a high hill, then comes speeding down, and makes a series of sharp turns during the descent. All around are brightly colored booths and other rides decorated with flashing lights. Carnival music is playing in the background as the roller coaster carries its screaming passengers along the winding course.	33	0.81 (0.97)	2.84 (2.82)	-1.69 (-1.22)
9-I. You are a champion racer at the Indianapolis 500 (where you go round and round the course). The competition is right up with you. The heat is very intense, and so are the gas fumes. You watch the dials and the road very carefully.	32	1.22 (1.71)	2.83 (2.45)	0.40 (2.54)

TABLE 4 (Cont'd)

Situation Description <sup>b</sup>	n	Pleasure	Arousal	Dominance
10-J. You are at the funeral of a distant relative. It takes place at the funeral home, and everyone is dressed in dark colors. The minister's voice drones on in the service. The casket up in front is the only decoration in the otherwise plain room.	27	-0.90 (-2.19)	-0.43 (-1.21)	-0.84 (-1.14)

<sup>a</sup>All numbers in parentheses are taken from Mehrabian and Russell (1974).

<sup>b</sup>Letters and numerals associated with situations are used in Figures 1A and 1B.

computed between the three emotional mediators and the 10 BI measures, across all snack situations (n=315). One significant canonical root emerged (chi square = 50.29, d.f. = 30, p < .025), showing a correlation of .33 between the criterion and predictor vectors. This relationship, while not particularly strong, is more substantial than under the ANOVA model where P,A and D mediators were not considered. In that case, the main effect for Situations explained only about one percent of the variance in BI; the canonical model indicates that subjects' reactions to situations can explain somewhat more of the variance. This lends further support to the notion of utilizing the psychological situation in examining situational influence.

Despite the apparent increase in explanatory power resulting from the use of the pleasure, arousal and dominance variables, it is evident that the situation, in and of itself, is not a powerful predictor of consumer behavior. This is to be expected. Considerable research has shown that personality variables consistently explain no more than 5-10% of the variance in consumer behavior (Kassarjian, 1971). The variance explained by situational variables most likely will be of the same order of magnitude. However, the interaction of situational, personality, and other classes of variables should lead to increased comprehension of consumption behavior.

#### Discussion

Earlier it had been stated that the two primary advantages of the Mehrabian-Russell framework were its potential for providing a partial explanation of situational influence and its taxonomical power. The latter claim received some support in the present research, which indicated that, across situations, subjects' responses demonstrated reliable differences along the three mediating dimensions of pleasure, arousal, and dominance (as evidenced in Figure 1A). If a large number of purchase situations were to be considered, the P,A, and D scores could be used in a clustering algorithm to allow classification of similar situations. This was not undertaken in the present case, however, due to the relatively small number of purchase situations being investigated.

The explanatory capabilities of the Mehrabian-Russell theory could not be fully investigated under the present research design. The three proposed mediating, explanatory dimensions can be regarded as generic dimensions of situations, which should explain generic situational influence. A set of ten snack product

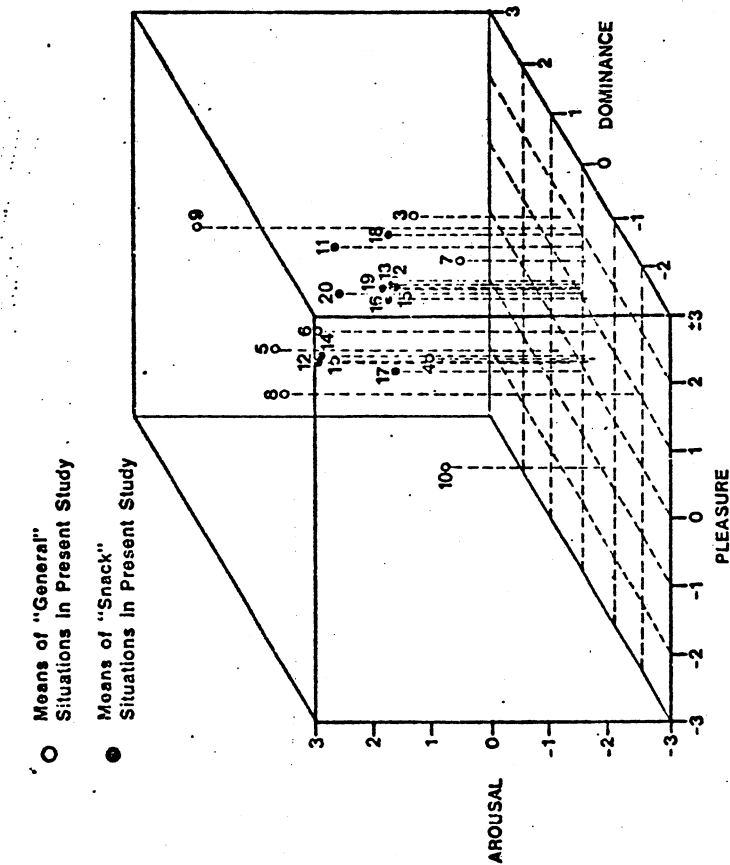


Figure 1A: Comparison of Group Means on Pleasure, Arousal and Dominance for Ten "General" Situations Selected from Mehrabian and Russell (1974).

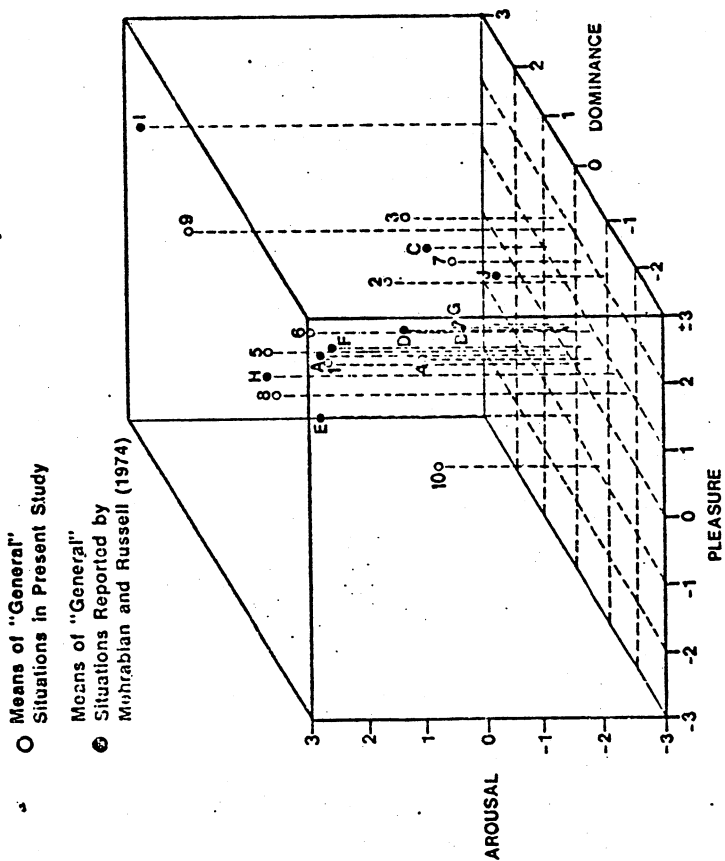


Figure 1B. Group Means on Pleasure, Arousal and Dominance for Ten "Snack" Situations (Belk, 1974) and Ten "General" Situations.

buying intentions does not meet this criterion. In future research, it would appear that decision processes and/or cognitive structure would be more appropriate dependent variables upon which to focus. Wright (1974) has recently called for the description of generic types of "decision environments" to aid in the prediction of the structure of consumer choice processes. Combination of generic situational dimensions with generic dependent variables representing choice processes will be a valuable area of research.

The specific dimensions of pleasure, arousal and dominance are also interesting because of their relationship to other trends in consumer research. For instance, the pleasure dimension may be quite useful in explaining context effects in attitude change situations (Razran, 1940; Janis, Kaye and Kirschner, 1965) or the effects of humor in advertising.

In Hansen's (1972) excellent treatment of the consumer choice process, he proposes that the discrepancy between "optimal" and "actual" levels of arousal is the primary determinant of the structure of the choice process. Further, arousal properties of the environment are conceived as the major impetus in the initiation of choice processes. Use of the Mehrabian-Russell framework to measure arousal may allow operationalization and testing of Hansen's theory, which offers the richest process orientation of the many "comprehensive" models of consumer behavior.

Finally, the dominance dimension of the Mehrabian-Russell theory closely parallels the work on "information overload" (Jacoby, et. al., 1974) and "overchoice" in the marketplace (Settle and Golden, 1974). In each of these situations, the consumer is supposedly intimidated by some aspect of his environment, causing him to engage in "dysfunctional" behavior. Similarly, research on internal/external locus of control would seem to benefit from consideration of the dominance dimension.

At a managerial level, situational research would appear to hold great potential for the construction of point-of-purchase displays. By pretesting to determine consumer's reactions along the pleasure, arousal and dominance dimensions, the optimum display for attracting customers could be identified.

Finally, the value of multiple approaches to the study of situational influence cannot be overstated. The Mehrabian-Russell framework may be useful for deriving generalizable statements at a generic level, but other approaches may be better suited to other objectives. Perhaps retail management could benefit from the application of Barker's (1968) "behavior settings" analysis. Similarly, Insel and Moos' (1974) approach to identifying specific situational dimensions may be quite useful in distinguishing among the various types of situations proposed by Hansen (1972). Regardless of the method employed, however, situational research must maintain an interactive focus by examining the relationship between situational and other variables. Otherwise the "situation" will quickly become one more addition to the already lengthy list of isolated constructs which are purported to, but do not, explain consumer behavior.

#### FOOTNOTES

1. This research was supported by an intramural grant from the Academic Senate of the University of California, Los Angeles, and by the Center for Marketing Studies, Graduate School of Management, UCLA.
2. Richard J. Lutz is Assistant Professor of Marketing, and Pradeep Kakkar is a

doctoral candidate in Marketing.

3. More appropriate data analytic strategies might view the Products dimension as a repeated measures factor, or as a set of dependent variables in a MANOVA design. The decision to treat it as a fixed factor was based on the desire for comparability of the present findings with Belk's earlier results.
4. In one sense, Belk's (1974b) procedure can be viewed as a series of experimental treatments (i.e., situation descriptions) followed by measures of behavioral intentions. Using Campbell and Stanley's (1963) notation, this would be diagrammed as  $X_1O_1X_2O_2X_3O_3 \dots X_{10}O_{10}$ . This type of design is especially susceptible to test-treatment interaction, which may partially explain the Product x Situation interaction effect.

#### REFERENCES

- Barker, R.G. Ecological psychology: Concepts and methods for studying the environment of human behavior. Palo Alto: Stanford University Press, 1968.
- Belk, R.W. Application and analysis of the behavioral differential inventory for assessing situational effects in buyer behavior. In S. Ward and P. Wright (Eds.), Advances in consumer research. Volume I. Urbana, Ill.: Association for Consumer Research, 1974(a), 370-80.
- Belk, R.W. An exploratory assessment of situational effects in buyer behavior. Journal of Marketing Research, 1974(b), 11, 156-163.
- Belk, R.W. Situational influence in consumer behavior. Unpublished manuscript, Temple University, 1974(c).
- Berlyne, D.E. Conflict, arousal and curiosity. New York: McGraw-Hill, 1960.
- Campbell, D.T. and J.C. Stanley. Experimental and quasi-experimental designs for research. In N.L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.
- Cottrell, L.S., Jr. Some neglected problems in social psychology. American Sociological Review, 1950, 15, 705-12.
- Frank, R.E., Massy, W.F., and Wind, Y. Market segmentation. Englewood Cliffs, New Jersey: Prentice-Hall, 1972.
- Gronhaug, K. Buying situation and buyer's information behavior. European Marketing Research Review, 1972, 7, 33-48.
- Hansen, F. Consumer choice behavior: A cognitive theory. New York: The Free Press, 1972.
- Hays, W.L. Statistics for psychologists. Chicago: Holt, Rinehart and Winston, 1964.
- Insel, P.M., and R.H. Moos. Psychological environments: Expanding the scope of human ecology. American Psychologist, 1974, 29, 179-88.



- Jacoby, J., Speller, D.E. and Kohn, C.A. Brand choice behavior as a function of information load. Journal of Marketing Research, 1974, 11, 63-9.
- Janis, I.L., Kaye, D. and Kirschner P. Facilitating effects of "eating-while-reading" on responsiveness to persuasive communications. Journal of Personality and Social Psychology, 1965, 1, 181-6.
- Kassarjian, H.H. Personality and consumer behavior: A review. Journal of Marketing Research, 1971, 8, 409-419.
- Mehrabian, A., and Russell, J.A. An approach to environmental psychology. Cambridge: M.I.T. Press, 1974.
- Moos, R.H. Conceptualizations of human environments. American Psychologist, 1973, 28, 652-65.
- Murray, H. Toward a classification of interaction. In T. Parsons and E.A. Shils (Eds.), Toward a general theory of action. Cambridge: Harvard University Press, 1952.
- Osgood, C.E., Suci, G.J., and Tannenbaum, P.H. The measurement of meaning. Urbana: University of Illinois Press, 1957.
- Razran, G.H.S. Conditioned response changes in rating and appraising sociopolitical slogans. Psychological Bulletin, 1940, 37, 481.
- Rotter, J.B. The role of the psychological situation in determining the direction of human behavior. In M.R. Jones (Ed.), Nebraska symposium on motivation. Lincoln: University of Nebraska Press, 1955, 245-69.
- Sandell, R.G. Effects of attitudinal and situational factors on reported choice behavior. Journal of Marketing Research, 1968, 5, 405-8.
- Settle, R.B., and L.L. Golden. Consumer perceptions: Overchoice in the marketplace. In S. Ward and P. Wright (Eds.), Advances in consumer research. Volume I. Urbana, Ill.: Association for Consumer Research, 1974, 29-37.
- Thomas, W.I. Primitive behavior. New York: McGraw-Hill, 1937.
- Toffler, A. Future shock. New York: Random House, 1970.
- Volkart, E.H. Introduction: Social behavior and the defined situation. In E.H. Volkart (Ed.), Social behavior and personality: Contributions of W.I. Thomas to theory and social research. New York: Social Science Research Council, 1951, 1-32.
- Wright, P.L. The harassed decision maker: Time pressure, distraction, and the use of evidence. Unpublished working paper, University of Illinois, 1973.
- Wright, P. The use of phased, noncompensatory strategies in decisions between multiattribute products. Unpublished working paper, Stanford University, 1974.

## Appendix A

## The Mehrabian-Russell Emotional Mediators (PAD) Instrument

Wide-awake	: +4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : -4 :	Sleepy	(A)
Controlled	: +4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : -4 :	Controlling	(D)
Melancholic	: -4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : +4 :	Contented	(P)
Aroused	: +4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : -4 :	Unaroused	(A)
Influential	: -4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : +4 :	Influenced	(D)
Awed	: +4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : -4 :	Important	(D)
Satisfied	: +4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : -4 :	Unsatisfied	(P)
Autonomous	: -4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : +4 :	Guided	(D)
Bored	: -4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : +4 :	Relaxed	(P)
Happy	: +4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : -4 :	Unhappy	(P)
Dominant	: -4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : +4 :	Submissive	(D)
Calm	: -4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : +4 :	Excited	(A)
Annoyed	: -4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : +4 :	Pleased	(P)
Hopeful	: +4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : -4 :	Despairing	(P)
Cared-for	: +4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : -4 :	In control	(D)
Relaxed	: -4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : +4 :	Stimulated	(A)
Dull	: -4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : +4 :	Jittery	(A)
Frenzied	: +4 : _____ : _____ : _____ : _____ : _____ : _____ : _____ : _____ : -4 :	Sluggish	(A)



## A SITUATIONAL MULTI-ATTRIBUTE ATTITUDE MODEL<sup>1</sup>

Kenneth E. Miller<sup>2</sup>  
University of Utah

Several researchers have discussed the promise of inclusion of situational variables in consumer research. Wicker (1969) states that measurement of attitudes and behavior should be carried out under similar situational conditions. This research investigates the variability of attitude scores when measured for differing situations and the resultant efficacy in the prediction of consumer preference and choice. The situational multi-attribute attitude model is utilized to predict situational preference and situational choice for each individual in the large mail panel. For these same individuals the nonsituational (traditional) model is utilized to predict nonsituational (overall) preference and nonsituational choice. The results from these sets of analyses are contrasted. The situational multi-attribute model, when used to predict situational choice, outperforms the traditional model, when used to predict overall choice. The success of situational multi-attribute model is necessarily product-specific, in this case the fast food hamburger market in Columbus, Ohio. However, the success of situational variables in improving prediction of brand choice for a specific product category certainly improves the promise of inclusion of these variables in consumer behavior.

### Introduction

A potentially profitable area for improving the ability of attitude toward a brand to predict choice of that brand over competing brands by a consumer is consideration of the situational variables which affect consumer behavior. Wicker (1969) states "a general postulate regarding situational influence on attitude-behavior relationships is the following: the more similar the situations in which verbal and overt behavioral responses are obtained, the stronger will be the attitude-behavior relationship."

There has been limited research reported in consumer behavior which considers the benefits derived from measuring attitude-behavior relationships along situational dimensions. Research indicates that situational variables have a significant influence on consumer affect towards a product category. Belk (1974) demonstrated that situational variables accounted for nearly half of the explained variance in both meat and snack preferences. Respondents preferred differing types of meat products in differing situations. Sandell (1968) found beverage preferences obtained for different situations, such as "for breakfast," "with lunch," "at a party," explained a larger proportion of total variance in reported behavior, than did favored brand. Thus, evidence

exists which suggests that explicit consideration of situational variables will improve the predictive ability of attitude and preference measures in the prediction of consumer choice.

The objective of this research is to compare the advantage of disaggregate v. aggregate situational analysis of attitude-behavior relationships for each individual in a specific product setting. The performance of a situational (disaggregate) model in predicting situational preference and choice will be compared with an aggregate model in predicting overall preference and choice for brands within a product category. The attitude model used is the multi-attribute model which has been widely used in marketing in both predictive and diagnostic applications (see Wilkie and Pessemier for an extensive review).

The general form of the model used in these studies is:

$$A_j = \left[ \sum_{i=1}^n V_i \left| B_{ij} - I_i \right|^k \right]^{1/K} \quad (1)$$

where:

$A_j$  = attitude toward brand  $j$

$V_i$  = affective importance of attribute  $i$

$B_{ij}$  = perceived amount of attribute  $i$  contained by brand  $j$

$I_i$  = ideal amount of attribute  $i$  (measured or assumed)

$n$  = number of attributes considered

$k$  = parameter of the weighted Minkowski space which determines the measure of distance calculations.

The axes of the perceptual space are the designated attributes for the product category and the attitude measure is represented as a measure of distance between the location of the brand and the product category ideal point in the multidimensional space. This model does not envisage situation-specific attitude measurement and is considered to be an aggregate model.

Situational measurement can be incorporated into the multi-attribute attitude model in several ways. It is hypothesized that perceived attribute importance will vary across differing situations. Depending on the specific product category chosen and salient attributes used, the performance of any brand along certain attributes also will vary across differing situations. The form of situational multi-attribute attitude model used is:

$$A_{js} = \left[ \sum_{i=1}^n V_{is} \left| B_{ijs} - I_{is} \right|^k \right]^{1/K} \quad (2)$$

where:

$s$  = situation (1,2---p)

$p$  = number of situations considered

$A_{js}$  = attitude toward brand  $j$  in situation  $s$

$V_{is}$  = importance of attribute  $i$  in situation  $s$

$B_{ijs}$  = perceived amount of attribute  $i$  contained by brand  $j$  in situation  $s$

$I_{is}$  = ideal amount of attribute  $i$  in situation  $s$

The sensitivity of  $V_{is}$ ,  $B_{ijs}$  and  $I_{is}$  across situations is dependent on product category and specific attributes considered.

### Hypotheses

- $H_1$ : The situational model does not perform better than the nonsituational model in the prediction of situational and overall preference, respectively.
- $H_2$ : The situational model does not perform better than the nonsituational model in the prediction of situational and overall choice, respectively.

### The Data

The data were collected from a mail panel residing in Columbus, Ohio. Each respondent completed five questionnaires, over a period of three months, concerning eight Columbus fast food restaurant chains. The initial panel sample of 744 was generated from a random list of names in the Columbus telephone directory.<sup>3</sup> During the three-month period an extensive advertising and couponing campaign was undertaken by one of these fast food restaurants.

Data collection was comprised of three stages:

- 1) Group interviews were conducted to isolate major situational usage patterns and salient product attributes;
- 2) A sample of 50 householders and 100 students were asked to indicate the frequency of visit(s) to fast food restaurants under each of the multitude of situations generated from (1);
- 3) Data were collected from the mail panel on the ratings of eight fast food restaurants on seven attributes (taste of food, speed of service, popularity with children, price, variety on menu, cleanliness, and convenience). The only brand rating obtained across situations was convenience. In addition, subjects were asked to rank preferences for the eight restaurants across situations and to rate the importance of each attribute for each situation. Choice information was also obtained from the panel. Thus, attribute importance, preference and choice as well as brand ratings on convenience, were obtained for all four situations considered.

The four situations chosen for detailed examination were selected with two criteria in mind. The situation should be encountered frequently so that adequate sample sizes could be obtained for each situation. Also, the situations should include different dimensions of situation. Belk (1974b) discusses five dimensions of a situation, being physical surroundings, time frame, interpersonal surroundings, mood and goal direction.

The four situations selected for detailed analysis were:

- to eat at lunchtime on a weekday
- to eat a snack after a shopping trip
- to eat when rushed for time
- to eat with the family when not rushed for time.

Lunch on a weekday evokes spatial and temporal dimensions for a major proportion of the panel.<sup>4</sup> Eating after a shopping trip evokes spatial and perhaps mood dimensions. Eating with the family when not rushed for time captures a consistent contrast in interpersonal surroundings. Eating when rushed for time encompasses time frame dimensions. Darley and Batson (1973) in their "Good Samaritan" experiment indicate that degree of haste in one's journey is a good situational predictor of helping behavior. The consumer's decision, as to which fast food restaurant to frequent, may be influenced by the amount of time he has available.

The respondents were also asked to complete an extensive AIO bank and media habit items. At the conclusion of the study, respondents were asked to give their opinions as to the nature of the study.

### Analysis

In order to ascertain the value of introducing the situational variables into the attitude-preference and attitude-behavior relationships, two sets of analyses were carried out for four time periods.

#### Attitude-Preference Analysis

For each individual who indicated frequenting fast food restaurants in at least two situations, the traditional model (model I) is contrasted to the predictive power of the situational model (model II). For each model, the seven attributes<sup>5</sup> were applied to each subject to compute an attitude score. An assumed ideal was used in both models (more of any attribute is better). The ranked attitude scores for models I and II were correlated with rank order overall and situational preference, respectively, using Spearman's Rho to determine the degree of association. This analysis is presented in Figure 1.

The mean Spearman's Rho across individuals for each situation at time periods 2, 3, 4 and 5 are presented in Table 1. The Wilcoxin matched sign test indicated that the magnitude of correlations were significantly different in only one out of 16 cases.

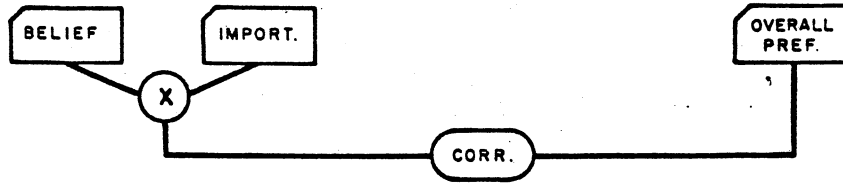
#### Attitude-Choice Analysis

A similar procedure was undertaken in the prediction of consumer choice. Following models I and II it is hypothesized that the brand closest to the product category ideal point will be the brand that is chosen. Situational and nonsituational attitude scores were computed and ranked. The brand closest to the assumed ideal point was ascertained and compared to the brand which was actually chosen by the respondent. The confusion matrix of brands predicted from the nonsituational attitude scores at period 1, given brand chosen from period 1 to period 2, is shown in Table 2.

The confusion matrix of brands predicted from the situational attitude scores at period 1, given brand chosen from period 1 to period 2, is shown in Table 3. This matrix is an aggregation of confusion matrices for all four situations.

Figure 1

AGGREGATE MODEL - Preference Correlation



SITUATIONAL MODEL - Preference Correlation

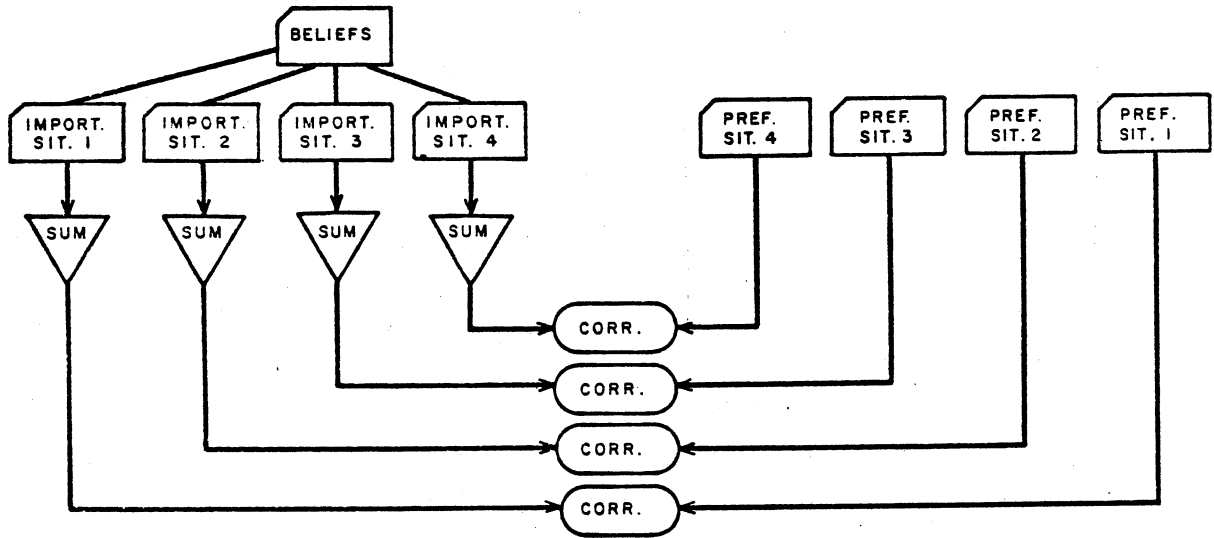


TABLE 1  
Comparison of Model I and Model II in the Prediction of Consumer Preference

		T <sub>2</sub>	T <sub>3</sub>	T <sub>4</sub>	T <sub>5</sub>
Lunch on a Weekday	I	.731 <sup>a</sup>	.764	.759	.779
	II	.722	.755	.743	.786
	n <sup>b</sup>	217	159	156	138
After a Shopping Trip	I	.738	.791	.776	.778
	II	.746	.783	.752	.812
	n	144	108	97	92
When Rushed for Time	I	.734	.784	.773	.823
	II	.748	.772	.761	.818
	n	205	154	153	128
With the Family When Not Rushed for Time	I	.752	.777	.778	.807 <sup>c</sup>
	II	.761	.782	.775	.781
	n	164	130	119	111

<sup>a</sup>Mean Spearman's Rho correlation coefficient for all subjects for each situation

<sup>b</sup>Sample size for specific situation. The total sample decreased during the study.

<sup>c</sup>Difference between models significant ( $p < .06$ )



TABLE 2  
 Confusion Matrix of Brand Predicted from the Nonsituational  
 Attitude Scores at Period 1 Given Brand Chosen During Period 1-2

		BRAND PREDICTED							
		A	B	C	D	E	F	G	H
BRAND CHOSEN	A	.177	.065	.0	.339	.0	.194	.113	.113
	B	.032	.258	.0	.168	.019	.303	.123	.097
	C	.035	.175	.140	.053	.0	.386	.158	.053
	D	.041	.061	.007	.456	.007	.299	.095	.034
	E	.100	.0	.0	.450	.0	.200	.150	.100
	F	.024	.048	.021	.125	.028	.561	.131	.062
	G	.033	.066	.007	.072	.013	.296	.414	.099
	H	.014	.057	.0	.071	.014	.257	.186	.400

% Correct = 39.81

TABLE 3  
 Confusion Matrix of Brand Predicted from the Situational  
 Attitude Scores at Period 1 Given Brand Chosen During Period 1-2

		BRAND PREDICTED								
		A	B	C	D	E	F	G	H	n
BRAND CHOSEN	A	.161	.065	.0	.339	.0	.194	.129	.113	58
	B	.026	.325	.039	.136	.039	.227	.104	.104	154
	C	.032	.063	.222	.079	.0	.349	.206	.048	63
	D	.020	.059	.0	.484	.0	.327	.085	.026	153
	E	.0	.0	.158	.368	.053	.105	.211	.105	19
	F	.021	.045	.021	.118	.007	.587	.149	.052	288
	G	.027	.048	.0	.089	.007	.253	.466	.110	146
	H	.014	.056	.0	.070	.014	.225	.099	.521	71

% Correct = 44.25

For each situational choice, the two models were used to predict this choice occasion, i.e., each choice was predicted separately. A matrix of brands predicted and brands chosen was constructed for each situation and the percentages of correct predictions are outlined in Table 4.

TABLE 4  
Comparison of Model I and Model II  
In The Prediction of Consumer Choice

		T <sub>1-2</sub> <sup>a</sup>	T <sub>2-3</sub>	T <sub>3-4</sub>	T <sub>4-5</sub>
Lunch on a Weekday	I	38.66 <sup>b</sup>	35.66	29.03	24.63
	II	45.25	44.66	41.17	32.87
After a Shopping Trip	I	30.21	35.00	27.77	35.71
	II	33.69	36.84	32.50	41.03
When Rushed for Time	I	43.68	50.36	42.98	37.23
	II	46.69	53.03	50.00	43.95
With the Family When Not Rushed for Time	I	41.87	46.15	41.13	44.94
	II	43.37	50.00	41.80	43.02
All Situations	I	39.81	41.00	34.75	32.64
	II	44.25	46.74	42.42	37.96
<sup>a</sup> Attitude scores at period 1 were used to predict choice reported at period 2. <sup>b</sup> Percentage correct prediction					

For the 20 comparisons of the model, the situational model outperformed the nonsituational model 19 out of 20 times. The Wilcoxin matched pairs test indicated that the magnitude of differences in choice prediction between the two-model forms was statistically significant ( $p < .001$ ).

#### Discussion

The results of the analysis illustrate that the disaggregate situational analysis is of value in improving the prediction of consumer choice, but is not superior to the aggregate model in the prediction of preference.

Several explanations may be advanced as to why the disaggregate analysis did not prove superior in the prediction of preference. Situational influence has been treated by certain consumer behavior theorists as mediating behavioral intention and choice.<sup>6</sup> It could be hypothesized that, for a single product category, preference judgments are without situational variance but choice decisions were influenced by predictable situations. The use of the rank ordered preference measure leads to the use of less powerful nonparametric statistical tests. If situation has a relevant but small influence on preference more powerful parametric statistical tests may be required to isolate the effect. Use of more rich measurement of preference will alleviate the problem.<sup>7</sup> However, consistent with the theoretical proposals for situation-specific attitude and behavior measurement, situational attitude yields significantly improved prediction of consumer choice. The value of inclusion of situational variables is necessarily product-specific. Various factors

favor the use of situation variables in the fast food hamburger restaurant market. The inclusion of attribute weights in the multi-attribute attitude model allow for variation of the number and nature of the salient attribute set across differing purchase situations. Also, the convenience of each restaurant varied with the spatial dimension of situation. These characteristics of the fast food consumer, in addition to the Columbus fast food restaurants exhibiting some degree of benefit segmentation, help explain why the inclusion of situation variables improves the relationship between attitude and behavior. Certainly the success of inclusion of situational variables in attitude measurement for brands in a product category, where attitudes, preferences and consumer choice are situation-specific, has ramifications for the consumer researcher.

Refinements to this research include examination of preference measurements to isolate situational variance. The power of situational rank order preference to predict choice should be contrasted to the predictive power of the overall preference measure. Analysis of variance should be conducted on attribute importance weights to ascertain situational variance. Examination of brand images along the specific situations considered in the research will give insight into the situational variation in attitudes for the brands. In addition to refinement and diagnosis of the research results presented here, the dimensions of situation need to be refined so that the study of consumers in predictable situational environments can be pursued.

#### FOOTNOTES

1. The research reported in this paper was supported in part by the College of Administrative Science, The Ohio State University. The author is indebted to Professor James L. Ginter for his constructive guidance during the research.
2. Kenneth E. Miller is Assistant Professor Marketing, College of Business, University of Utah.
3. The final panel size, reduced by experimental mortality, was 446.
4. E.g., those who were employed away from their place of residence.
5. The popularity with children attribute was not applied to all respondents.
6. E.g., Sheth (1969).
7. Pessemier, et al. (1971) present a dollarmetric preference measure which is interval scaled.

#### REFERENCES

- Belk, R. W. An exploratory assessment of situational effects in buyer behavior. Journal of Marketing Research, 1974, 11, 156-163.
- Belk, R. W. Situational influence in consumer behavior. Faculty Working Paper No. 195, College of Commerce and Business Administration. University of Illinois at Urbana-Champaign, 1974.

- Darley, J. M. & Batson, C. D. From Jerusalem to Jericho, a study of situational and dispositional variables in helping behavior. Journal of Personality and Social Psychology, 1973, 27, 100-108.
- Fishbein, M. Attitude and prediction of behavior. In M. Fishbein, Readings in attitude theory and measurement. New York: John Wiley and Sons, Inc. 1967, 479-92.
- Ginter, J. L. A study of attitude change and choice behavior during new product introduction in a laboratory setting. Unpublished doctoral dissertation, Purdue University, 1972.
- Ginter, J. L. An experimental investigation of attitude change and choice of a new brand. Journal of Marketing Research, 1974, 11, 30-40.
- Pessemier, E. A., Burger, P., Teach, R. & Tigert, D. Using laboratory brand preference scales to predict consumer brand choice. Management Science, 1971, 17, B371-B385.
- Sandell, R. G. Effects of attitudinal and situational factors on reported choice behavior. Journal of Marketing Research, 1968, 5, 405-08.
- Sheth, J. N. Attitude as a function of evaluative beliefs. Paper presented at the American Marketing Association Workshop in Consumer Behavior, 1969.
- Wicker, A. W. Attitudes versus actions: the relationships of verbal and overt behavioral responses to attitude objects. Journal of Social Issues, 1969, 25, 41-78.
- Wilkie, W. L. & Pessemier, E. A. Issues in marketing's use of multi-attribute attitude models. Journal of Marketing Research, 1973, 10, 428-41.



ATTRIBUTION THEORY IN MARKETING RESEARCH:  
PROBLEMS AND PROSPECTS

R. E. Burnkrant<sup>1</sup>  
University of California, Berkeley

Attribution theories address the issue of how people infer, from limited available evidence, unobservable attributes or dispositions about the objects and organisms in their environment. As such, they are theories about how people go beyond the directly observable "data" to infer further elements, that is, to complete a partial representation of some focal object. Attribution theories, then, are theories about how people make attributions. They attempt to account for the conditions in which and the extent to which people are able to infer dispositions or states in other organisms or objects from limited available evidence.

The bulk of the attribution theory research may be traced back to Heider's (1958) work on causal inference in which he addressed the issue of how people attribute causes to the events or occurrences in their environment. This work was extended by Jones and Davis' (1965) theory of correspondent inferences and Kelley's (1967) presentation of attribution theory. Both of the latter contributions sought to further specify the conditions under which observers are able to infer unobservable dispositions or attributes in the organisms or objects they encounter. These two contributions have had a direct impact on the consumer behavior applications of attribution theory.

Attribution Research in Marketing

The bulk of the marketing research on attribution theory published to date has focused on communications issues. This work, which we might lump together as representing the Settle school, has sought to specify the conditions under which a person's statements about an object will tend to be informative about that object.

In a study entitled "Attribution Theory and the Acceptance of Information," Settle (1972) found that the type of source which would be most informative about an object depends on the type of object about which information is sought. His independent variables were product type (i.e., complex, visible, durable, or multipurpose) and source type (i.e., expert, close friend, own experience over time, or own experience in a variety of ways). His dependent variable was the subject's confidence in a "good" product choice. No attempt was made to deal with, explicate, or measure any part of an attribution process. The only relevance of the study at all to attribution theory is in the rather loose correspondence of his source types to three of Kelley's inference validating criteria (i.e., consistency over time, consistency over modality, and consensus).

In one of this group's earliest studies, Settle, Faricy, and Warren (1971) manipulated the consistency of the responses actors made to each of several movies. By varying this consistency the researchers found that subjects were able to predict how a movie would be rated if the movie were con-

sistently evaluated by several actors and, conversely, how an actor would rate a fourth movie if he rated all previous movies consistently and uniquely. They also measured the subject's confidence in these ratings.

It was suggested that people make attributions about others on the basis of the others' consistency of responding. Implicit in this is the assumption that, if a person gives the same response to several objects and if this response is unique to that person in the sense that it is not also made by others, some sort of biasing disposition such as the motivation to sell the objects is attributed to the person. On the other hand, if he makes a distinct or different response to each object and if these responses are consistent with those made by others the person would not tend to be seen as having some biasing disposition; and, hence, his responses should be informative about the object being evaluated. In this latter instance the individual would be seen as an objective observer and what he says would tend to be believed.

This early study led to the prediction that a two-sided message may lead to more attitude change in the recommended direction than a one-sided message. Settle and Golden (1974) attempted to provide support for this conclusion by exposing subjects to messages which were either one-sided or two-sided. They then measured the degree of confidence the subjects had in each statement that the focal brand was superior to brand X on a specific attribute. They combined the confidence rating with the importance the individual assigned to that attribute to obtain a derived measure which they called "expected value." They interpret their results as supporting the proposition that a two-sided message is more effective than a one-sided message.

They suggest that subjects in the one-sided condition attribute to the source the motivation to "sell his particular brand." Presumably then, subjects in the two-sided condition attribute to the source the motivation to objectively inform or provide information about the brands. It is suggested that these attributions affect the believability of the message and, hence, the resultant belief scores about the focal object.

It seems to me that there are substantial problems with this research. The first problem concerns their failure to measure the attributions made by the subjects or to provide other validating evidence to support their contention that their treatments manipulated these attributions. The obtained dependent measures (i.e., predicted ratings or "expected values") were *assumed to be the result of attributions made by the subject*. These attributions, however, were not verified by any direct measurement. They were merely assumed to mediate the independent-dependent relationship under consideration. In the absence of substantiating evidence their contention remains merely an interesting hypothesis or *possible* explanation for the obtained results.

Attributions have frequently been measured in the literature. For instance, Calder and Burnkrant (1973) asked subjects to describe an actor on a set of bi-polar adjective scales which yielded three personality dimensions. McArthur (1972) asked subjects to assign the cause of an actor's behavior by indicating whether it was due to something about the *person*, something about the *stimulus object*, something about the *particular circumstances*, or some combination of these causes. It is heartening in this regard to see that Mizerski (1974) attempted to provide a somewhat similar measure in his study when he asked subjects to estimate the extent to which "other reasons--reasons having nothing to do with" the stimulus object, affected the actor's opinion.

A second and somewhat related problem is that, in the absence of direct measures of the attribution process, the contribution to our understanding of communications processes is minimal. They deal with the same variables treated in the extensive Yale Communications Research Program (e.g., Hovland, Janis, and Kelley, 1953) although they propose an attributional explanation for the observed phenomena. Thus, the finding that a two-sided message is more effective than a one-sided message is not all that unique. In their classic study in this area Hovland, Lumsdaine, and Sheffield (1949) found that, with an educated audience (by which they meant high school graduates), a two-sided message is more effective than a one-sided message at inducing belief change in the recommended direction. Since the college students used by Settle and Golden (1974) are clearly consistent with what Hovland, Lumsdaine, and Sheffield consider to be an audience high in education their finding is not very surprising.

#### Prospects for Future Research

While an attempt to actually measure the attributions people make about a message source and the effects these attributions have on attitude change certainly seems warranted, there appear to be other marketing oriented issues to which attribution theories can offer a more unique contribution.

Since people make attributions about others in all situations every day as a normal part of coping with their environment, these attributions seem likely to exert an important influence on the purchase decisions people make. Therefore, research which addresses itself more directly to the verification and further specification of the inference process and its role or roles in consumer behavior seems to be badly needed.

There, no doubt, are many directions which this research could take. In the following material I would like to specify some areas to which I think attribution theories can contribute by looking at research in which I have been involved and some of the research questions related to this.

One approach which seems to merit investigation is to focus on the extent to which people base their purchase decisions on the attributions others are expected to make about them on the basis of their use of particular product brands in specific situations. In this regard Calder and Burnkrant (1973) found that observers are able to attribute personality dispositions to others on the basis of their use of particular brands in a given social context. They found, further, that the attributions made by an observer are dependent upon the social context in which the brand is used.

This study may be viewed as the first step in a research program directed at uncovering the role played by attributions about users in product purchase decisions. In other words, it may be that buyers and users of products are aware that others will make attributions about them on the basis of their product usage in particular situations. This knowledge may influence their own purchase decisions. It may be, for instance, that people buy products and use them in particular situations in order to reinforce or vary the concept or perception others have about them. This seems consistent with the literature which has attempted to relate the self-concept to the perception people have of the typical user of their favorite brand.

If these conjectures are supported by research evidence it would then seem appropriate to try to specify the conditions, both in terms of situational



and product dimensions, under which these attributions would be most important in determining purchase. One may hypothesize, for instance, that the expected attributions of others would be important determiners of purchase where the individual plans to use a conspicuous product in a highly involving situation.

Attribution theory seems also to have considerable potential for integrating and adding meaning to the substantial body of literature on product perception and social influence. One of the important criteria of any good theory, however, is that it leads us to ask more insightful questions capable of advancing our knowledge; and, in this regard too, attribution theory seems relevant to this literature.

Commonly, an individual in a purchasing context is faced with the problem of evaluating the worth of a product which he cannot possibly objectively evaluate from physical manipulation and observation of its characteristics alone. In these situations the individual would be expected to turn to other items of information which he has learned to use as a basis for inferring the characteristics of the object. Among the cues that have been found effective in various situations are price (e.g., Valenzi and Andrews, 1971), brand (e.g., Allison and Uhl, 1964), and the behavior of others (e.g., Venkatesan, 1966).

Therefore, we might expect these cues to affect the individual's perception of the product to the extent that the individual is unable to directly observe and objectively evaluate the product through the use of cues intrinsic to the product. It should, perhaps, be noted that Olson and Jacoby (1972) found that what they call "intrinsic cues" were seen as better predictors of quality than extrinsic cues such as price.

Attribution theory indicates that the confidence in an inference made about an object would be stronger if the attributional implications of all cues associated with the object were consistent than if they were inconsistent. Further, the theory would suggest that a cue would be discounted to the extent that reasons other than those related to the characteristics of the product could have accounted for the cue's presence. Both derivations seem to be directly applicable to the area of product perception.

Many other applications of attribution theory could, no doubt, be suggested. These applications, however, should consider, in much more direct terms than has been recognized in the previously published applications of the Settle school, the extent to which inferences mediate aspects of (buyer) behavior, the types of inferences which mediate this behavior, and the conditions under which these inferences may be reliably obtained. These, after all, are the considerations which provide the central focus and unique contribution of the attribution theories.

#### FOOTNOTE

1. R. E. Burnkrant is Assistant Professor, Schools of Business Administration, University of California, Berkeley.

## REFERENCES

- Allison, R. I., and Uhl, K. P. Influence of beer brand identification on taste perception. Journal of Marketing Research, 1964, 1, 36-39.
- Calder, B. J., and Burnkrant, R. E. Attributions based on consumer behavior. University of Illinois, College of Commerce and Business Administration. Working paper number 130, 1973.
- Heider, F. The psychology of interpersonal relations. New York: Wiley, 1958.
- Hovland, C. I., Lumsdaine, A. A., and Sheffield, F. D. Experiments on mass communication. Princeton: Princeton University Press, 1949.
- Jones, E. E., and Davis, K. E. From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), Advances in experimental social psychology. Vol. II, New York: Academic Press, 1965.
- Kelley, H. H. Attribution in social psychology. Nebraska Symposium on Motivation, 1967, 15, 192-238.
- Mizerski, R. W. A test of the relationship between trait and causal attribution. Paper presented at the Fifth Annual Conference of the Association for Consumer Research, Chicago, November, 1974.
- Olson, J. C., and Jacoby, J. Cue utilization in the quality perception process. Paper presented at the Third Annual Conference of the Association for Consumer Research, Chicago, November 1972.
- Settle, R. B. Attribution theory and acceptance of information. Journal of Marketing Research, 1972, 9, 85-88.
- Settle, R. B., Faricy, J. H., and Warren, G. T. Consumer information processing: Attributing effects to causes. Paper presented at the Second Annual Conference of the Association for Consumer Research, College Park, Maryland, September 1971.
- Settle, R. B., and Golden, L. L. Attribution theory and advertiser credibility. Journal of Marketing Research, 1974, 11, 181-185.
- Valenzi, E. R., and Andrews, I. R. Effect of price information on product quality ratings. Journal of Applied Psychology, 1971, 55, 87-91.
- Venkatesan, M. Experimental study of consumer behavior, conformity and independence. Journal of Marketing Research, 1966, 3, 384-387.



A TEST OF THE RELATIONSHIP  
BETWEEN TRAIT AND CAUSAL  
ATTRIBUTION

Richard W. Mizerski<sup>1</sup>  
Arizona State University

Attribution principles, developed to predict trait attribution in person perception, were applied to the causal attribution of information about products in order to test an attribution theory explanation for unfavorable information's disproportionate influence in consumer decision-making. Subjects received either favorable or unfavorable ratings about one of two products. It was hypothesized that subjects who received unfavorable information would make more extreme internal attributions (perceive the information as "caused" by the product only), as well as have more confidence in their attributions. The results show that the predicted attributions appeared for one of the two products, with no significant differences in attributional confidence. Probable causes and implications of the findings are discussed.

In this era of consumer protection, marketers are being faced with an unparalleled barrage of unfavorable information about their products or services. Consumer-oriented publications (e.g., Consumer Reports), counter advertising by organizations such as Stern Concern, comparative advertising by competitors, corrective advertising, and widely publicized tests of products by government agencies (e.g., E. P. A. gas consumption tests of new automobiles), are just a few of the sources that have augmented word-of-mouth advertising as potent media for unfavorable product information. While the chances of a business encountering potentially ruinous and possibly untrue information are increasing, the marketing literature has just begun to seriously analyze its impact.

The limited marketing literature that deals with unfavorable information suggests that it wields an inordinate influence in an individual's decision-making process. Arndt (1968) reported that only 18% of the subjects in his study purchased a test product after they received unfavorable word-of-mouth communication; yet 54% of those who received favorable word-of-mouth communication purchased the same item. Reynolds and Darden (1972) elicited responses from women who rejected a fashion innovation (the midi clothing style), and found that 93% perceived their discussions about the fashion as unfavorable. Although neither study controlled for the amount of each type of information, they did provide empirical data on unfavorable information's differential impact.

Social psychologists have provided more rigorous analyses of the perceptions of favorable and unfavorable information in their studies of impression formation. This research (e.g., Anderson, 1965; Feldman, 1966) typically has subjects evaluate a stimulus person based upon manipulated sets of equally polarized favorable and unfavorable adjectives. Their findings substantiate those of Arndt in marketing, and show the disproportionate influence of unfavorable information (adjectives) in forming overall impressions.

An understanding of the differential strength of favorable (usually the firm's communications) and unfavorable information could help marketers gauge the total effect of communications about their products, and ultimately improve the promotion strategy of their companies. Unfortunately, attempts to explain this perceived differential in strength have not been successful. A number of reasons such as unfavorable information's greater surprisingness (Feldman, 1966), or its less frequent use (Zajonc, 1968), have been proposed. However, these reasons have been almost totally eliminated by Kanouse and Hanson (1971).

Another possible explanation is based upon the concept that individuals may differ in their processing of favorable and unfavorable information, and that the difference centers on the source or "cause" to which the information is attributed. This area of information processing is referred to as attribution theory, and involves "an examination of the process whereby people attribute characteristics, intentions, feelings, and traits to objects or events in the real world" (Kanouse and Hanson, 1971). This is accomplished by an analysis of phenomenal causality; i.e., noting the determinants and consequences of causes for particular actions or events (Jones and Davis, 1965). In terms of processing information, attribution theory suggests that an individual accepts and is influenced by information from another person only if he attributes the report to the entity being described ("internal" cause). Therefore, an accurate report about an entity is one that is perceived to be "caused" by that entity. In many cases, however, an individual may believe that the information was "caused" by another factor or factors such as the possible social pressure to flatter an individual (one possible "external" cause).

One principle of attribution theory is that "the role of a given cause in producing a given effect is discounted if other plausible causes are also present" (Kelly, 1973, p. 113). This concept, often referred to as the "discounting principle," has been demonstrated in experiments of compliant behavior (Thibaut and Riecken, 1955), and in the perceived sincerity of complimentary ratings (Jones, Gergen, and Jones, 1963). Both studies suggested that their subjects perceived more external causes for favorable behavior or information (causes other than the entity on which information was being provided), and thus the possibility of the information having an internal (originating in the entity) cause should be discounted. This discounting supposedly leads to a stronger inference of external causation. These results led Kelly to propose that "complimentary remarks made about a person under the guise of accurately reporting one's feelings about him . . . may be attributed to some other causes such as the ambient social situation" (1971, p. 18). By applying the discounting principle, favorable information should have a greater tendency (than unfavorable information) to be attributed to external causes.

The principle of "social desirability of effects" (Jones and Davis, 1965) provides a possible explanation for the attribution of unfavorable information. Viewed as a modification of the discounting principle (Kelly, 1973), social desirability of effects has two assumptions:

1. In general, people intend desirable effects from their actions.
2. Socially desirable effects provide relatively little information about causes. When a person adheres to social norms (e.g., intending desirable effects), one is unsure whether the person's behavior was caused by internal (true intentions, beliefs) or external (society, peer pressure) causes. Therefore, socially undesirable behavior tends to rule out external causes, and leaves internal causes as the reason(s) for the effect (Hastorf, et al, 1970).

Kanouse and Hanson (1971) have provided an interpretation and extension of this rationale. They suggest that because of social norms of discussions about people or objects, most information is favorable in an individual's social environment. Therefore, unfavorable information about traits or behavior would stand out and have a tendency to be attributed to the person or object being evaluated (internal cause). An example of this attribution behavior would be the following:

If an individual believes that most people are sincere, he is likely to be relatively unimpressed with the information that person x is sincere. Given that sincerity is the norm, credit attaches to social pressure or to the simple fact of being human rather than to person x. Insincerity, however, is another matter. By standing in contrast to the norm, the insincere individual invites attributions of responsibility for this trait. (p.57)

Both of the attribution principles, social desirability of effects and the discounting principle, have evolved from studies dealing with the attribution of traits in person perception. Their application depends upon perceived social norms of discussion about individuals. It would seem that the perception of products (based upon information about their attributes) involves an analogous process, one in which the principles of attribution should apply if there are similar perceived norms of discussions about products. A possible problem in generalizing these concepts from a person to a product domain is that the attribution of a trait has always been inferred rather than measured directly. This raises a question as to whether there would be significant differences in attribution if the subjects were asked to actually specify the cause for the information. In order to test the appropriateness of generalizing the concepts, the following hypothesis was developed:

Subjects will tend to attribute favorable information about a product to "external" causes, and unfavorable information about a product to "internal" causes.

If the differences in attribution do exist for processing favorable and unfavorable product information, this differential should affect information influence. An individual would seemingly give more weight to information he felt reflected a product's characteristics (the suggested attribution of unfavorable information), than to information that tends to be attributed to

external influences such as social norms (the suggested attribution of favorable information).

Another facet of attribution theory may augment a theoretical explanation of unfavorable information's differential influence. Jones and Davis (1965) proposed that more confident and extreme attributions follow from observed effects that deviate from the norm (the suggested perception of unfavorable information). They argue that confidence and extremity of trait attribution were "equivalent and inseparable" measures. In other words, one can make highly confident trait attributions only when the stimulus person does something distinctive, thus also justifying an extreme trait attribution. If this is the case, one would logically expect that individuals would tend to impute a greater weight to information in which they had more confidence. The following hypothesis will test the role of confidence in the causal attribution of information about products:

Subjects will have more confidence in causal attributions about unfavorable information than with causal attributions about favorable information.

#### Experiment 1 Methodology and Results

In order to determine if there seem to be social norms about product discussions, and thus some consensus about the type of information (favorable or unfavorable) expected, 97 subjects<sup>2</sup> were asked whether they would expect "favorable" or "unfavorable" information from: (1) a student who had just tested a new automobile, and (2) a student who had just previewed a new movie.<sup>3</sup> The responses to these questions (Table 1) show that a majority of subjects did expect favorable information about each product (78% for the automobile, 59% for the movie). Since similar tests in trait attribution have not been conducted,

TABLE 1  
Expected Information Type<sup>a</sup>

Expected Information	Automobile	Movie
Favorable (+)	78%	59%
Unfavorable (-)	18%	23%
Undecided	4%	19%

<sup>a</sup>n = 97 subjects

there are no previous measures from which to determine whether either product would have the minimum expectancy necessary in order to apply the proposed attribution principles. However, a simple majority does suggest a feasible application.

Experiment 2  
Methodology

The main experiment was performed on 9 classes (300 subjects)<sup>4</sup> of upper division undergraduate and graduate students at the University of Florida College of Business Administration. Each student was told that the study 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, method) of a product to which they were randomly assigned. The automobile attributes were maintenance costs, comfort and gas mileage. Acting, plot and the photography were the three movie attributes. The information was presented in the form of personal ratings, supposedly reprinted from 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. The modifiers used for rating were scaled for equal polarity and affective opposite meaning by replicating a study by Myers and Warner (1968). Following the information treatment, subjects were asked to respond to an attribution question similar to the following:

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 <u>opinion</u>
	1	2	3	4	5	6	7
Maintenance costs:							

The smaller the scale value chosen, the more internal ("caused" by the product only) the attribution. After the attribution question, the respondents were asked to rate their confidence in each attribution. A seven point scale ranging from "no confidence" (#1) to "complete confidence" (#7), recorded their confidence scores.

Results

The first hypothesis was supported for the movie only. Mean attribution scores for the two products are presented in Table 2. While more extreme

TABLE 2  
Mean Attribution Scores

Treatment	Automobile	Movie	Combined
Favorable (+)	3.40	3.56	3.48
Unfavorable (-)	3.52	2.90	3.21
Combined	3.46	3.23	



internal attributions (smaller mean scores) were hypothesized to occur in response to unfavorable information for both products, only the movie scores are in the predicted direction. In addition, unfavorable information produced more extreme internal attributions across products. An unweighted means analysis of variance (ANOVA) was used in order to detect significant differences between treatments (Table 3), and revealed a main effect of attribute ( $F = 7.46$ ,  $p < .001$ ), and an interaction of product and information ( $F = 4.33$ ,  $p < .038$ ). The main effect of attribute shows that the subjects perceived a difference

TABLE 3  
Analysis of Variance: Mean Attribution Scores

Source	df	MS	F
Attribute (A)	4	17.75	7.46 <sup>a</sup>
Info treatment (I)	1	4.91	2.06
Product (P)	1	3.42	1.44
A x I	4	3.04	1.28
I x P	1	10.30	4.33 <sup>b</sup>
Error	258	2.38	

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

<sup>b</sup> $p < .038$

between the attributes in terms of each being the cause for their individual bogus ratings. However, these differences were independent of the information type received and do not constitute a test of the hypothesis. In order to detect the source of the information x product interaction, a Newman-Keuls Multiple Range Test was performed (Table 4). The results of this test show that

TABLE 4  
Results of A Newman-Keuls Multiple Range Test on  
Mean Attribution Scores

Information	Comparison Groups			
	Unfavorable	Favorable	Unfavorable	Favorable
Product	Movie	Auto	Auto	Movie
Means	2.90	3.40	3.52	3.56
	-----			

----- Not significantly different at  $p < .05$

only the movie's attribution scores were significantly different in the two information conditions, with unfavorable information producing the predicted stronger internal attributions.

The second hypothesis was not supported for either product. In fact, the mean confidence scores (Table 5) show favorable information producing greater confidence, a result opposite that predicted. However, an unequal means ANOVA

TABLE 5  
Mean Confidence Scores

Treatment	Automobile	Movie	Combined
Favorable (+)	4.97	4.65	4.81
Unfavorable (-)	4.70	4.44	4.57
Combined	4.83	4.55	

failed to find any significant differences between the group means, which suggests that the differences are due to chance.

While the predicted significant differences in confidence failed to appear, it is still of some interest to determine whether the concept of "correspondence" is generalizable to the causal attribution of information about products. A series of polynomial regression analyses were applied to the data in order to measure the degree of association between the extremity of causal attribution, and the confidence in that attribution. A high degree of association, particularly for the movie, would lend support to the Jones and Davis position of "inseparability" (i.e., if the same principles are the basis for each phenomenon, their reactions to stimulus should be highly related). Results of the regressions (Table 6) reveals that three of the four information x product conditions show significant relationships. However, the number and mix of non-linear components, as well as the inconsistency of their appearance across conditions, reveal that the relationship is very complex. On balance, it does appear that causal attribution and confidence in the attribution are related, but the idea that they are "equivalent and inseparable" (as suggested in trait attribution), is not supported by the data.

#### Discussion

It is interesting to note that the movie, not the automobile, conformed to the predictions of attribution theory. Since more subjects expected favorable information about the automobile than about the movie (Table 1), the automobile would seem to be the more obvious product for application of the attribution principles. Perhaps the different degree of objectivity of judgment for the two product's attributes accounts for the hypothesis being supported for the movie only. The movie's attributes called for more subjective evaluation than two of the three automobile attributes. It may be that the more objective the criteria, the less the perceived difference in causal attribution. Subjects who were interviewed after the experiment seemed to think that attributes such

as maintenance costs and gas mileage involved rather straight-forward ratings based upon "hard" empirical data, which would leave little room for "other" (external) causes to effect the evaluation. On the other hand, evaluating attributes such as acting skill, the plot, and the quality of the photography depend on more subjective criteria (external causes) such as past experience and learned tastes. This suggests that the principles developed in trait attribution may be particularly applicable in situations where subjective evaluations are important<sup>5</sup> such as with many products using social, psychological, or any emotional appeal (e.g., see Settle and Mizerski, 1973). Future research should include some measure of the subjects' perception of the ability to objectively evaluate a product's attributes so that this possibility could be tested.

TABLE 6  
Polynomial Regression: Attribution on Confidence  
in The Attribution

Degree Polynomial	Automobile (+) <sup>a</sup>		Automobile (-) <sup>b</sup>	
	F-Value	Explained Variance <sup>e</sup>	F-Value	Explained Variance
1	0.55	---	0.22	---
2	2.00	---	12.88 <sup>f</sup>	17.6%
3	0.47	---	3.58	---
4	2.95	---	1.53	---
5	0.65	---	0.20	---
	Movie (+) <sup>c</sup>		Movie (-) <sup>d</sup>	
1	0.90	---	2.66	---
2	6.12 <sup>g</sup>	7.1%	2.32	---
3	4.31 <sup>h</sup>	5.0%	0.55	---
4	4.35 <sup>h</sup>	5.1%	6.06 <sup>g</sup>	8.2%
5	2.24	---	0.17	---

<sup>a</sup>n=67, <sup>b</sup>n=61, <sup>c</sup>n=74, <sup>d</sup>n=68, <sup>e</sup>Explained variance is approx.

<sup>f</sup>p<.001

<sup>g</sup>p<.025

<sup>h</sup>p<.05

The failure of negative information to evoke more confident attributions may have been a function of the very limited stimulus information provided the subjects. Many respondents noted that their confidence in the attribution would have been substantially altered if they had been given more information about the bogus rater. Of course, the object of the study was to eliminate those factors and test only for the effect of attribute and information. However, future studies could manipulate information about the source (e.g., product experience, or demographic profile), or even the source of the message (government, independent testing organization, consumer group, etc.) in order to give subjects a more realistic information matrix. These manipulations would also be interesting in terms of the resulting causal attributions.

Finally, the tested application of two principles of trait attribution was partially successful, and does suggest at least part of an explanation for the disproportionate influence of unfavorable information. Future research could tie these and other principles of trait attribution to established models of decision-making (e.g., Mizerski, 1974) in order to provide a more rigorous analysis of the concept of causal attribution in consumer purchase decisions.

#### FOOTNOTES

1. Richard W. Mizerski, Assistant Professor of Advertising and Marketing, Arizona State University.
2. The subjects were similar to those used in the main (#2) experiment.
3. These are the same products used in the main experiment.
4. For the analysis, 30 subjects were deleted for missing data and for improper use of a scale used in research external to this paper. The latter deletion was necessary because improper scale use affected the data collected for this study.
5. Perhaps the success in applying these principles in trait attribution stems from those studies' test subjects' perception of the generally subjective nature of evaluating others.

#### REFERENCES

- Anderson, Norman H. Averaging versus adding as a stimulus--combination rule in impression formation. Journal of Personality and Social Psychology, 1965, 2, 1-9.
- Arndt, Johan. Perceived risk, sociometric integration, and word-of-mouth in the adaption of a new food product. In D. F. Cox (Ed), Risk taking and information handling in consumer behavior. Boston: Division of Research, Graduate School of Business Administration, Harvard University, 1968.
- Feldman, Sheldon. Motivational aspects of attitudinal elements and their place in cognitive interaction. In Sheldon Feldman (Ed), Cognitive consistency. New York: Academic Press, 1966.
- Fishbein, M. and Rayen, B. The AB scales: An operational definition of belief and attitude. In M. Fishbein (Ed), Readings in attitude theory and measurement. New York: John Wiley and Sons, Inc., 1962.

- Hastorf, A.H., Schneider, D.J. and Polefka, J. Person perception. Reading, Mass.: Addison Wesley, 1970.
- Jones, E.E., Gergen, K.J. and Jones, R.G. Tactics of integration among leaders and subordinates in a status hierarchy. Psychological Monographs, 1963, 77.
- Jones, E.E., and Davis, K.E. 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.
- Kanouse, D.E., and Hanson, L.R. Jr. Negativity in evaluations. In Edward E. Jones, et. al. (Eds.), Attribution: Perceiving the causes of behavior. Morristown, N.J.: General Learning Press, 1971.
- Kelly, H.H. Attribution in social interaction. In Edward E. Jones, et. al. (Eds.), Attribution: Perceiving the causes of behavior. Morristown, N.J.: General Learning Press, 1971.
- Kelly, H.H. The processes of causal attribution. American Psychologist, 1973, 28, 107-128.
- Mizerski, R.W. Attribution theory and consumer processing of unfavorable information about products. Unpublished doctoral dissertation, University of Florida, 1974.
- Myers, J.H., and Warner, G. Semantic properties of selected evaluation adjectives. Journal of Marketing Research, 1968, 5, 409-412.
- Reynolds, F.D., and Darden, W.R. Why the midi failed. Journal of Advertising Research, 1972, 12, 39-46.
- Settle, R.B., and Mizerski, R.W. Differential response to objective and social information in advertisements. In Conceptual and methodological foundations of marketing. (Proceedings of the 1973 Fall A.M.A. conference.) Chicago: American Marketing Association, 1974.
- Thibaut, J.W., and Riecken, H.W. Some determinants and consequences of the social causality. Journal of Personality, 1955, 24, 113-133.
- Zajonc, R.B. Attitudinal effects of mere exposure. Journal of Personality and Social Psychology, 1968, 7, 1-29.

COMPOSITE POPULATION DESCRIPTORS: <sup>1</sup>  
THE SOCIO-ECONOMIC/LIFE CYCLE GRID

Richard B. Ellis<sup>2</sup>  
American Telephone & Telegraph Co.

Greater sophistication in marketing data and analysis techniques has brought increased and more precise insight, but has also caused problems in maintaining consistency between studies and communicating results to management. Two composite population descriptors - Socio-Economic Status and Family Life Cycle - have been found valuable both as surrogates for some more complex factors and as stabilizing parameters in presentations and comparisons of study findings.

Both measures have been updated and adapted to the specific needs of the communications market, and their validity and efficiency in this work is described. Combined in a matrix format, the "SES/LC Grid", they have proven useful in the fields of defining strata for survey probes, correlation and extrapolation of results from different studies and normalization of different populations for comparison purposes.

The past several years have seen an almost exponential increase in the diversity and sophistication of the data and analytic techniques available to the student of consumer behavior. The benefits of this inventive growth are apparent to anyone who has followed their exposition in the literature, but it has also produced a few impediments for the researcher whose goals include the development of a body of consistent analytic findings over time for management applications. Specifically, three problems have tended to accrue in this environment:

- While increasing sophistication has improved our insight and understanding, its complexity has also increased the problems of making these findings clear and credible to the client of research - the manager without professional analytic skills.
- In observing consumer behavior over time, it is obviously of benefit to be able to link and measure its changes, but this becomes quite difficult as the definitions and measurements of its dimensions shift and alter as new developments appear.
- Finally, when studying different manifestations of market effects on a geographic basis, there is a natural desire to anchor these in some way for comparative purposes, so that differences caused by location, product perception and sales effort may be identified and evaluated.

None of these difficulties preclude the use of advances in the art, but they do indicate to the prudent practitioner the need for some research yardstick against which can be measured the wealth of new data we are acquiring. The criteria for such a yardstick are practically self-evident. It should be:

- Simple. Easily comprehended and grasped.
- Logical. Possessing prima facie validity.
- Consistent. Durable enough to withstand both the winds of change and the fires of fad.
- Comparable. Capable of being linked not only to other internal information, but external data as well.

The most satisfactory tool we have found to forestall these difficulties, given these criteria, is the use of two composite demographic measures in combination - Socio-Economic Status and Family Life Cycle measurement.

These measures have been so familiar to research practitioners for a number of years that they will bear only the briefest review here. However, a few words in defense of these sometimes maligned measures may be in order. It is undoubtedly true that the actual coalescence of vague intentions into an act of purchase is an exceedingly complex process, fraught with stimuli and responses pyramided on values and attitudes nurtured deep in the consumer's bosom. But it is also true that many of these relatively ephemeral factors are at least partially mirrored and defined in overt descriptive facts about the buyer, and that empirically strong correlations can be found between buyers' characteristics and their consumption patterns.

At a time when research tends to center on the shifting values of social class and the consumption patterns of the post-nuclear family, this may seem a strange position. But the fact remains that our experience in research done on the residence telephone market strongly supports this contention. The composite variables themselves have been used extensively as inputs in a variety of marketing problems and analytic methodologies. The consistency and strength with which they emerge as significant descriptive and differentiating factors has confirmed their validity for this type of work in our market.

As one example, in a study using factor analysis, designed to identify the significant elements differentiating customer groups in terms of geography and purchase behavior, only four significant factors were found. Two of these primary descriptors were Socio-Economic Status and Family Life Cycle which, when coupled with family-housing and mobility factors accounted for 88% of the variability found among 86 population groups tested in the country.

It is also quite relevant that in several analyses, one or more of the components of these composite variables has been found to be significant on its own, and has entered the calculations in addition to the combined values. This has tended to verify the basic hypothesis that the combination of the individual components in the composite form apparently describes customer attributes different from those delineated by the individual variables alone.

Based on this experience, we believe that there is substantial evidence that these descriptors reflect certain relatively intangible life style and social factors which affect buyer behavior. Naturally, they should not be understood to be such factors themselves, nor are they directly substitutable for them. However, they have been found to correlate highly with these more complex measures, and they are considerably easier and less expensive to determine, analyze and understand. In addition, they have the advantage of being factual descriptors of our customers which can be used to estimate total population characteristics.

One final observation on these problems is germane, but more specific to the communications product than some others. Both of these descriptors obviously tend to measure the joint characteristics of families or households, in most cases these being synonymous. They have proven to be unusually efficient in our work since, unlike a number of other consumer products, the household, rather than a particular member, is the actual "consumer" of our services.

Socio-Economic Status (SES) is a numerical expression, developed by the U.S. Bureau of the Census, of the combined effects of three characteristics of the household and its nominal representative, the head of household. The three factors are income, education and occupational status. Each of these characteristics has been fitted to a numerical scale running 0-100, based on the household population of the country. The SES score is computed by determining the numerical value for each individual attribute and then averaging these three figures. For those interested in a more detailed description of the history and development of these factors, the Bureau of the Census has documented its methodology and findings quite thoroughly.

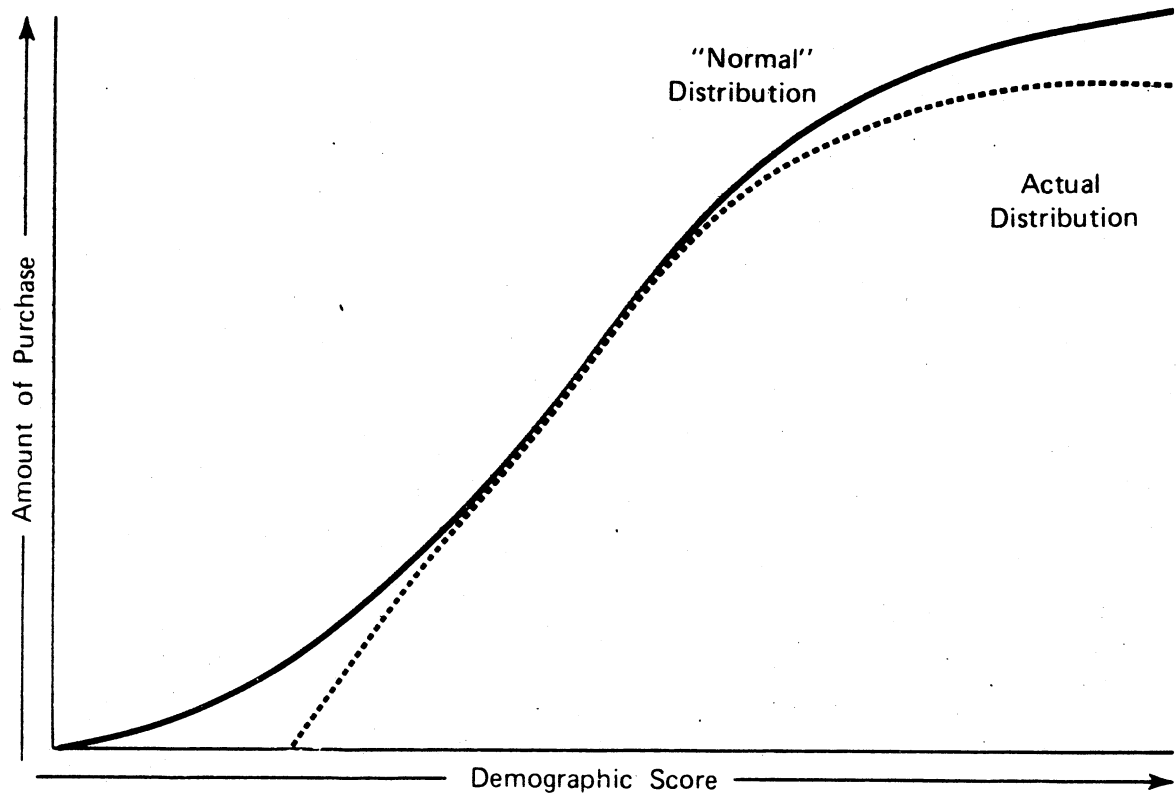


FIGURE 1. TYPICAL DISTRIBUTIONS OF SES VARIABLES VS. CONSUMPTION



In the original formulation, both the individual scores and the composite SES scores were selected so that the total U.S. population fell approximately on a normal distribution curve. When applied to market analysis of consumption patterns, of course, this assumption of normality does not hold strictly true. In the case of purchase behavior toward discretionary telephone products and services, the curve is truncated in the lower portion by an economic and cultural threshold below which customers do not perceive anything other than basic telephone service to be of value. It is also flattened in the upper portion where a simple extrapolation of ability and proclivity to purchase a product far outstrip the capacity to consume it. However, in the middle range of the distribution where the mass of the market lies, the basic assumptions have been found to pertain, and clear correlations between customer characteristics and buying behavior exist.

These concepts are illustrated in Figure 1, which generally reflects in prototype form the distribution of the SES component variables and the composite values as well. The assumed normal distribution in cumulative form is shown as a solid line, and a typical consumption pattern for communications products as a dotted line.

Since each component of the SES score has a discrete and limited number of possible values in the real world, the selection of the intervals for each individual variable and the composite score was done by a simple empirical fitting process. That is, purchase patterns for each product and service were determined for every possible value of each variable, and the break points established where the significant differences in behavior could be found. While this technique might seem peculiarly simple and naive, it proved quite effective, and a high degree of consistency was found among the individual variables and the composite scores related to the specific purchase patterns being studied.

Tables 1 and 2 show the original Census values assigned to family income and head of household education, and the specific scores selected for use in telephone market work. The differences are due to two basic reasons:

- The Bureau of the Census scores were computed and validated based on the population enumeration of 1960. Since that time, both average educational attainment and income have risen substantially due to various social and economic factors. This has tended to make the very fine gradations in the lower reaches of these tables less meaningful in terms of differentiating behavioral categories.
- A number of studies of the Bell System residence market have shown that there are certain relatively broad categories of income and education classification within which no significant differences among customers can be found.

For example, from a telephone usage viewpoint it is impossible to differentiate between households with incomes of \$6,000 and \$7,000, or household heads who left college after one year or two years. Therefore, no valid purpose is served by maintaining the more complex scoring system.

**TABLE 1**  
**SES Income Scores**

Bureau of the Census		AT&T	
Income Category	Score	Income Category	Score
\$ 3,000 - 3,499	21	Under \$3,000	15
\$ 3,500 - 3,999	27	\$ 3,000 - 4,999	31
\$ 4,000 - 4,499	34		
\$ 4,500 - 4,999	41		
\$ 5,000 - 5,499	49		
\$ 5,500 - 5,999	57	\$ 5,000 - 7,499	62
\$ 6,000 - 6,499	63		
\$ 6,500 - 6,999	69		
\$ 7,000 - 7,499	74		
\$ 7,500 - 7,999	78	\$ 7,500 - 9,999	84
\$ 8,000 - 8,499	81		
\$ 8,500 - 8,999	84		
\$ 9,000 - 9,499	87		
\$ 9,500 - 9,999	89	\$10,000 - 14,999	94
\$10,000 - 14,999	94		
\$15,000 - 24,999	98	\$15,000 - 19,999	97
\$25,000 or more	100	\$20,000 - 29,999	99
		\$30,000 or more	100

**TABLE 2**  
**SES Education Scores**

Bureau of the Census		AT&T	
Grade Completed	Score	Education Attainment	Score
Elementary - None	1	Some Grade School	10
1 - 2	2		
3 - 4	4		
5 - 6	8		
7	13		
8	23	Grade School Graduate	23
High School	1	Some High School	42
	2		
	3		
	4		
College	1	High School Graduate	67
	2	Some College	86
	3		
	4		
	5 or more	98	College Graduate
		Any Graduate School	98

**TABLE 3**  
**SES Occupation Scores**

Occupational Category	Score
Professional, technical and kindred workers	90
Managers, officials and proprietors (except farm)	81
Clerical, sales and kindred workers	71
Craftsmen, foremen and kindred workers	58
Operators and kindred workers	45
Service workers, including private households	34
Laborers, except farm and mine	20
Housewives	33
Students	33
Retired	33
Unemployed	33

In the case of occupational codes, the Bureau of the Census offers a very detailed set of scores reflecting close to 450 different occupations, plus a set of seven summary categories customarily used in most census reports. Similar to the cases of income and education, the very detailed classifications were found not to be particularly meaningful in differentiating telephone customers, so the summary categories were used, supplemented by some additional classifications to take into account "unemployed" but consuming households. These are shown in Table 3.

As a result of these analyses, two categorizations of SES were defined for use in our research projects, as shown in Table 4, together with the original Census classifications for comparison.

**TABLE 4**  
**Socio-Economic Status Categories**

Detail Classifications		Summary Classifications	
<u>Census</u>	<u>AT&amp;T</u>	<u>Census</u>	<u>AT&amp;T</u>
0-9	0-24	0-19	0-44
10-19	25-44		(Lower)
20-29	45-59	20-49	45-69
30-39	60-64		(Lower Middle)
40-49	65-69	50-79	70-89
50-59	70-74		(Upper Middle)
60-69	75-79	80-99	90-99
70-79	80-84		(Upper)
80-89	85-89		
90-99	90-99		

The ten interval "detail" category is designed to be used as an independent variable in a variety of technical methodologies, most often in conjunction with the specific variables which make up the composite SES score. It should be noted again that the classification intervals are not distributed evenly over the assumed normal curve, but rather are defined in specific terms of communication market patterns, generally concentrated in the middle ranges where the market for these services exists.

The four "summary" categories are derived from these classifications, and are used as interpretive criteria in presentations, sample stratification and population normalization, as described later in this paper.

The definitions and classifications of Family Life Cycle (FLC) status are less clear in terms of number of segments and significant characteristics. The demographer's viewpoint tends to center on key factors affecting population size and make-up. Probably the bestknown expositions of this facet are the articles by Glick and Parke, while Norton offers an updated view of the significant shifts in family life patterns. These have overall value in forecasting gross markets or economies, but they tend to be less useful to the market analyst faced with the problems of specific products in a relatively short-term environment.

Certain marketing applications are patently self-evident - more babies will eat more baby foods, marriages will produce housing starts, and an increasing older population will create demand for geriatric products. But for the vast majority of consumer products, the relationships tend to be more complex. General studies of FLC status related to broad patterns of income and consumption have produced some consistency on the significant components, age and make-up of family being the two most common. However, no consensus on the number and definition of the prototype groupings has emerged.

Wells and Gubar offer a comprehensive summary of these problems and describe the findings of various life cycle studies in the decade preceding 1966. They also propose a prototype classification, without specific values, which served as the basis of our work.

Faced with this uncertainty and definitional vagueness, the analyst is constrained to explore the subject anew, once again with specific products, services and consumption patterns in mind. Such a study was conducted for all the discretionary purchase activity in the residence telephone market utilizing all the variables suggested in the literature as potential components of an FLC measure. These included number, relationship, age and sex of family members, and an employment variable for older families as suggested by the 1964 Michigan Survey of Consumer Finances.

Sonquist and Morgan's Automatic Interaction Detector technique was selected for these analyses because of the variety and intercorrelation of the variables. The results were analyzed to produce a set of classifications significant both for the individual products and services, and the total patterns of expenditure.

TABLE 5

## Family Life Cycle Status Categories

Prototype (Wells & Gubar)	Detail Classifications				AT&T	Summary Classifications
	HOH Age	Fam. Size	Children < 12	> 12	Employ. Status	
Bachelor Stage	< 55	1	-	-	-	Younger No Children
Newly Married	< 55	2	-	-	-	
Full Nest I (Younger)						Younger Young Children
1+child < 6	< 55	3+	1+	0	-	
0 Child < 6	< 55	3+	-	1	-	Younger Older Children
Full Nest II (Older)	< 55	3+	-	2+	-	
Empty Nest Employed						Older
Retired	> 55	-	-	-	Emp.	
Solitary Survivor Employed	> 55	-	-	-	Not Emp.	
Retired						

Note: "--" indicates inapplicable variable.

Table 5 shows the general prototype classifications and the specific categories, with their definitions developed for telephone market analysis. The intended applications of the "detail" and "summary" classifications are identical to those defined for the SES categories.

Several interesting points emerged in the analysis which developed these classifications.

- The head of household age which discriminated most successfully between "younger" and "older" households came out to be 55. This is quite different from the traditional 65 year retirement, and breaks at the 40-45 year level developed in some previous studies, most notably the 1957 Life Study and the 1964 Michigan Survey already cited.
- The significant age which discriminated between "young" and "older" children was also different from figures generally used in previous studies. The breaking point for our purposes came at 12 years, rather than at 6 as suggested by a number of other analyses. The 1965 Life

Study did propose 12 as a discriminating age, but only in a secondary sense. This older age is undoubtedly due to the consumption patterns of telephone services as contrasted to a number of other consumer goods.

- In the case of "older" households (head of household over 55) the size and makeup of the household became irrelevant, and the employment status was the sole discriminating variable found.

The final synthesis of these composite descriptors for application purposes takes the form of the simplest analytic tool of all, the matrix. The format of this "Socio-Economic/Life Cycle Grid" is shown in Figure 2. As a matter of interest, the percent of U.S. households served by the Bell System, as estimated from our sample, is shown for each cell.

Several features should be noted in this arrangement:

- The four SES categories appear in the columns of the grid, and the four FLC categories as rows. Each cell, therefore, represents a specific and discrete class of people completely defined by the two factors. This cross classification ensures that the significant differences between the cell members - within cell variability - is minimized.
- For each SES and FLC category, the composite value of the characteristic displayed is aggregated for the entire class of customers in the marginal boxes, and the total value for the entire group is shown in the lower right hand corner.
- Certain key composite classes of customers are computed in the boxes at the bottom of the page - "Middle Class" (Lower Middle plus Upper Middle SES), "Family Households" (Younger Households with any children), "Younger Households", and "Middle Class Family Households" (the center four cells in the matrix).

In research applications, the SES/LC Grid is used in different ways depending on the complexity of the characteristics being studied, and the size of the sample being distributed through the 16 cells. For large sample populations, individual cell analysis of average characteristics may be valid for the aggregate or composite totals. When smaller populations are involved, the aggregate cells may be used for quantitative analysis, and judgements made on individual cell behavior dependent on the consistency of the data and the stability of relationships between the individual cells and their corresponding aggregates.

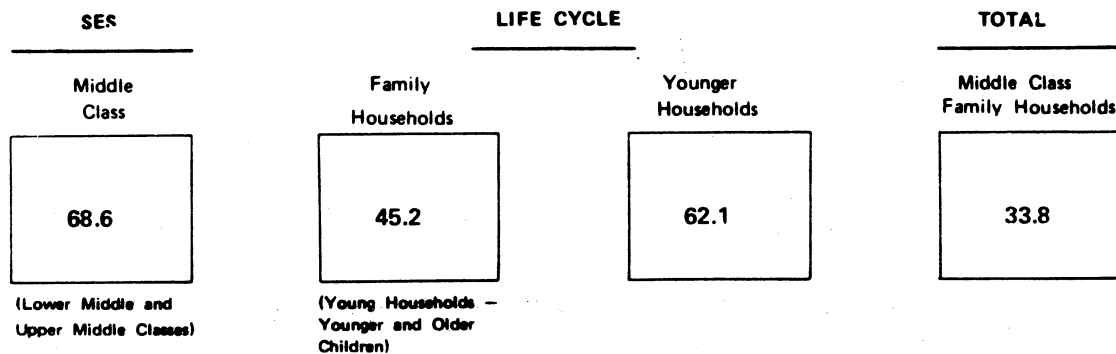
In practice, the grid has been used in three major applications, all somewhat interrelated - stratified sample selection, inter-project study correlation, and population normalization. The first application is relatively simple, but is quite valuable in terms of time and cost savings

**SOCIO-ECONOMIC/LIFE CYCLE GRID**

Population \_\_\_\_\_ Characteristic \_\_\_\_\_

Life Cycle Category		Socio-Economic Class				Life Cycle Totals
		Lower	Lower Middle	Upper Middle	Upper	
Younger Households	No Children	1.5	5.7	6.8	2.9	16.9
	Younger Children	0.9	6.2	8.1	4.3	19.5
	Older Children	1.4	8.7	10.8	4.8	25.7
Older Households		12.5	13.5	8.8	3.1	37.9
SES Class Totals		16.3	34.1	34.5	15.1	100.0

**COMPOSITE TOTALS**



**FIGURE 2**

for certain types of surveys. Typically, the situation involves preliminary probes on new concepts, products, services, pricing plans, or the like. The researcher is interested in refining and quantifying his hypotheses on a regional or national basis, but the selection of a probability or scientific sample would involve very large numbers of respondents, high cost and a longer time to complete and analyze the survey. This condition is substantially aggravated when visual materials or product demonstration are required to elicit a suitable respondent reaction, and in-home interview costs rise rapidly.

The normal answer to this dilemma is a stratified sample allowing for smaller numbers of respondents and more limited interviewing areas. However, lacking more definitive data on the segments of the new market configuration being studied, the strata are difficult to select and the size of the sample problematical. We have had considerable success in establishing the distribution of the study population in terms of the SES/LC Grid and using the 16 cells as the de facto strata. This allows much greater freedom in selection of interview techniques and locations; for example, a small number of fixed walk-in facilities strategically located in various regions of the country. A few basic demographic questions added to the study specific material categorize the respondents with regard to the cells in the grid. Interviews are continued until adequate representation is obtained in the cells, and the results analyzed. For quantified estimates of total population response, these results can be weighted by the known cell sizes and expanded mathematically.

While obviously not as accurate and precise as a carefully selected representative sample, it has proven quite valuable and less expensive in studies during the early stages of product or concept development.

A correlary of this technique is the ability to compare the results of studies conducted for different purposes, under different conditions, at different times and places. This is of particular concern in the communications market where a high degree of interactivity has been found between customer perceptions and purchases of different products and services. This same relationship appears to carry forward, as expected, into new products as well. The inclusion on each survey vehicle of the few responses necessary to classify the respondents allows certain basic comparisons to be made between various studies in terms of the estimated customer groups and their characteristics.

For example, in one set of preliminary studies on three new products, it was found during post-survey comparisons, that the supposed market for the different products overlapped by almost two-thirds. The desirability and cost of a composite of the three products should obviously be determined before proceeding with final design specifications.

Normalization of data is a standard and frequently used statistical technique. However, when the data being normalized are people, not only do methodological problems crop up, but a faint odor of manipulation may tend to taint the findings in the manager's mind. The objective is normally quite clear - the equalization of external factors so that a valid judgment may be made of the significant differences between two or more populations with regard to an identified characteristic or activity. The difficulty encountered with consumer groups involves the large number of exogenous factors which may pertain, and the complexity of trying to account for all of them.



The SES/LC Grid offers a simple visual method of accomplishing this, with the further advantage of the client's being able to see precisely what is being done and what the effects are. In actual practice, two techniques are used to accomplish this. In the case of large sample populations, (over 600) cell by cell analysis is generally possible and will yield a relatively precise picture of the legitimate differences between the two populations. Lacking this number of cases to study, the simplest approach is to proportionalize one of the populations to the structure of the other, and conduct the analysis on the resulting marginals. This is normally most easily accomplished by recomputing the marginal values for one population using the actual cases by cell of the other, thus eliminating the distortion caused by the disparities in distribution between the two.

As a byproduct of such comparisons, a general estimate of market potential is also practicable. Given two market groups broken into the SES/LC cells, a comparison of market development on a cell by cell basis will give clear indications of those population segments where purchase is distinctly less in one of the markets. The difference between the two will normally yield a rough estimate of the unrealized potential in the underdeveloped market.

The use of life style and social class data is generally accepted in consumer research today, and they offer unusual and valuable insights into the market and consumer behavior. However, it is also difficult and expensive to gather valid data and apply them to the total population. Use of Socio-Economic Status and Family Life Cycle factors, individually or in combination, offers a much less expensive and more stable alternative for many purposes. As a supplement to simple demographic characteristics, composite or index variables hold great potential for understanding and predicting buying behavior, extending our view without shortening our communications lines to our clients.

#### FOOTNOTES

1. This research was sponsored by the Market Research Section of the American Telephone & Telegraph Company. The author is indebted to Mr. G. Williams for substantial systems and data processing support in this project.
2. Richard B. Ellis is Marketing Manager - Reports of the Market Research Information System at AT&T.

#### REFERENCES

- Barton, S.G. The life cycle and buying patterns. In Lincoln H. Clark (Ed.), Consumer behavior. Volume II. New York: New York University Press, 1955.
- Glick, P.C. The life cycle of the family. Marriage and Family Living, February, 1955.
- Glick, P.C. & Parke, R., Jr. New approaches in studying the life cycle of the family. Demography. 1965, 2.

- Lansing, J.B. & Kish, L. Family life cycle as an independent variable. American Sociological Review, October, 1957.
- Lansing, J.B. & Morgan, J.N. Consumer finances over the life cycle. In Lincoln H. Clark (Ed.), Consumer behavior. Volume II. New York: New York University Press, 1955.
- Norton, A.J. The family life cycle updated: Components and uses. Paper presented at the annual meeting of the Population Association of America, New Orleans, La., April, 1973.
- Sonquist, J.A. & Morgan, J.N. The detection of interaction effects. University of Michigan, Institute for Social Research, Survey Research Center. Monograph No. 35, 1964.
- U.S. Bureau of the Census. Methodology and scores of socioeconomic status. Working Paper No. 15, 1963.
- U.S. Bureau of the Census. Socioeconomic characteristics of the population: 1960. Current Population Reports, Series P-23, No. 12, 1964.
- Wells, W.D. & Gubar, G. Life cycle concept in marketing research. Journal of Marketing Research. 1966, 3.



AWARENESS AS AN INDICATOR OF NEW PRODUCT PERFORMANCE<sup>1</sup>

David W. Olson<sup>2</sup>  
 Leo Burnett U.S.A.

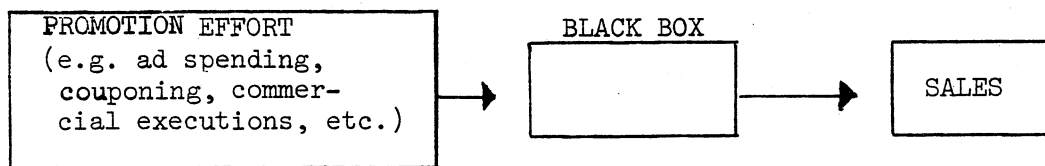
Since advertising weight for new products does not directly relate to trial levels achieved by those products, awareness is postulated to be a useful intervening variable between advertising weight and trial. Aided brand awareness is found to be a superior measure to unaided brand awareness in this application. At the same time, aided brand awareness is subject to problems of "false" awareness and of bias due to question order, and is a generally unstable measurement. Therefore, care must be taken to avoid designing models whose output is highly sensitive to varying inputs of awareness.

Traditionally, an integral part of any research package measuring performance of a new product in a test market is the A-T-U (Awareness-Trial-Usage) study. Awareness of the new product is generally considered to be one relevant indicator of performance for a new product. And yet, the precise nature of these variables rarely seems to be fully understood by those using awareness to evaluate the new product's performance.

The purpose of this paper is to clarify the nature of the measurement of awareness of new products, to specify the usefulness of such measures, and to caution against their misuse in certain applications.

## The Model

The basic form of various models which simulate new product performance is:



There is, of course, no simple relationship between promotion effort and sales for a new product. That is, what happens in the "black box" has a major effect on sales levels for any particular new product, which far outweighs the unitary effect of promotion effort.

The first step in expanding the model is, of course, to break "sales" into its components of trial and repeat purchase (usage). It is quite well accepted that repeat purchase is primarily a function of satisfaction with the product by triers, and that promotion effort only plays a secondary role in generating repeat purchase. Therefore, the most that can be hoped for is some relationship between promotion effort and trial levels achieved by new products.

Let the model be further restricted by considering one important aspect

of promotion effort, namely a measure of advertising weight in television.<sup>3</sup> Does there exist a relationship between TV advertising weight and trial levels achieved by new products?<sup>4</sup>

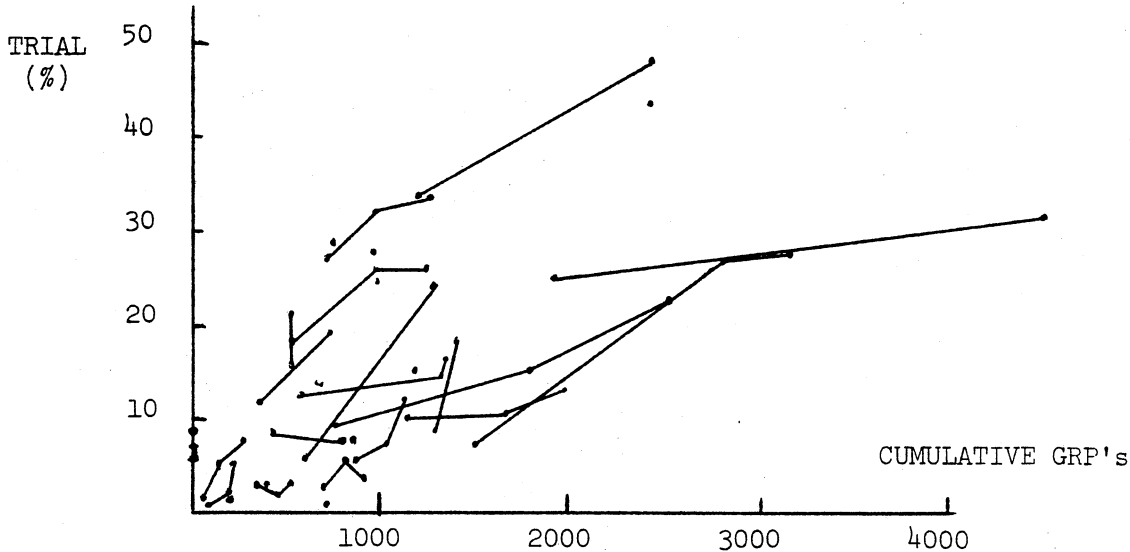


Fig. 1. Relation between cumulative Gross Rating Points on TV and trial for new products among potential market in first three months.

As Figure 1 shows, little variability in trial levels for different new products is accounted for by a measure of advertising weight (measured in terms of cumulative TV Gross Rating Points since introduction of the new product, confined to its first three months). The explanation is obvious: too many important factors in addition to ad weight (e.g. price, distribution, appeal of product concept) significantly affect trial for the relationship to hold.

The model is, therefore, of no predictive or quantitative use in this form. What is needed to improve it is some intervening variable, which relates well to both ad weight and trial, independently. Some measure of awareness of the new product seems a likely candidate. Intuitively, one should expect awareness to be significantly affected by advertising weight, and similarly, one can postulate that awareness relates, in some sense, to trial (if only as a necessary prerequisite to trial). The model would then look like this:



It is important to note that awareness, used in this way, is not a behavioral measure. It is nothing more than an artificial variable which, it is hoped, relates quantitatively to television GRP's and to trial. The balance of this paper will explore "awareness" as it relates as an intervening variable within the simple model described above.

#### Measurements Of Awareness

Awareness is far too often thought of as part of some sort of bipolar process, where consumers are either "aware" or are not. It is assumed that after repeated exposure to some stimulus, such as advertising, at some

point consumers move into a "state of awareness" from the state of total ignorance they were in previously.

Unfortunately, any thought on the matter at all will soon lead to the conclusion that awareness is, in fact, a continuum. In terms of a new product, awareness of that product may build from a dim recognition of some of the narrative elements in its commercial, through to brand name recognition, through to a detailed understanding of what the product is and how it may be personally beneficial. Section III will discuss a frequently employed method of incorporating awareness into models of consumer behavior which, in essence, violates the assumption that awareness is a continuum.

Given that awareness is a continuum, how does one measure it? One must understand that whatever they are measuring they are, by nature of their measuring tool, artificially dividing the continuous world into those "aware" and those "not aware."

Two measurements of awareness which have traditionally been obtained in studies are unaided and aided brand awareness. These two measurements will be discussed below, and evaluated on the basis of how each works as an intervening variable in the model which has been described.

#### Unaided Brand Awareness

Unaided brand awareness is the traditional measure of awareness for established brands. Countless studies have shown this measurement to be sensitive to changes in advertising weight for ongoing brands (e.g. Palda, 1969) and other studies have shown some relations between unaided awareness and purchase behavior (e.g. see Axelrod 1968, Assael and Day, 1968, Gruber 1969). It is, therefore, reasonable to explore how well it relates to advertising spending for new products.

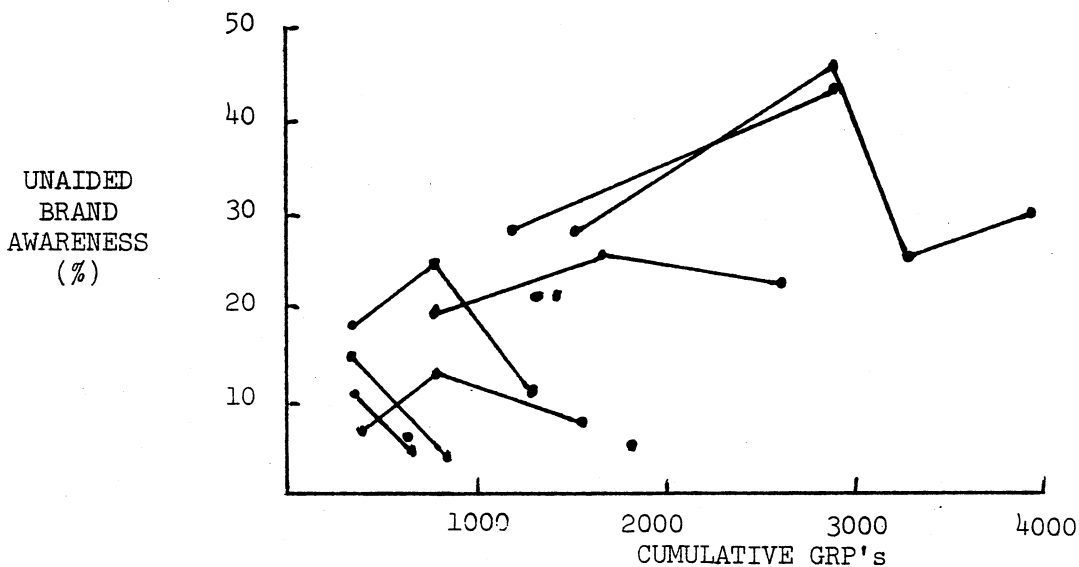


Fig. 2. Relation between cumulative Gross Rating Points on TV and unaided brand awareness for eight new cereals.

Figure 2 represents the relationship between cumulative GRP's since introduction and unaided brand awareness achieved by eight new cereals. As can be seen, the relationship is poor. It appears that the unaided awareness level obtained by a new cereal is more dependent on other factors than on the advertising weight spent behind that new product, and hence is not a good measure to incorporate within the model as it has been outlined.

There are three major reasons why the measurement of unaided awareness is unsuitable for new products as an intervening variable. First, new products generally attain rather low levels of unaided awareness, which makes discrimination between different brand performances somewhat difficult. Second, the numbers obtained are sensitive to how the question is asked. If asked to name all brands of cereal they can think of, a significantly smaller percentage of respondents will name a particular new cereal than if they were asked to name all the new cereals they can think of. Standardization of the question, especially across categories, becomes a hazardous task.<sup>5</sup> Finally, unaided awareness is probably a composite measure of a combination of other performance factors, such as recency of trial or usage. That is, unaided awareness may, to a sizeable extent, be generated by events which occur after trial of the new product, which is contrary to the implied causal structure of the model.

For these reasons, then, unaided awareness appears to be a poor measurement of awareness for new products to use in the model. Unaided awareness is not a useful intervening variable in this context.

#### Aided Brand Awareness

Aided brand awareness ("Have you ever heard of \_\_\_\_\_?"), to a large extent, overcomes the three problems of unaided awareness. It can be asked consistently, across product categories, and for products which have no relevant category; the magnitude of the numbers obtained are sufficiently large to provide a possibility for discrimination; and it is not unduly affected by events which occur after trial of the product.<sup>6</sup> At the same time, it has other problems, to be discussed shortly.

But first, the question is: Does aided brand awareness relate to advertising spending? The answer is yes; at least, it relates far better than unaided brand awareness.

Figure 3 is the relation between cumulative GRP's spent behind products and the aided awareness levels attained by them in the first three months after introduction. While the relation is by no means clean, the curve can be reasonably represented by an exponential curve

$$Y_t = a + (1-a)(1 - e^{-bS_t}) \quad 0 < a < 1 \quad (1)$$

where

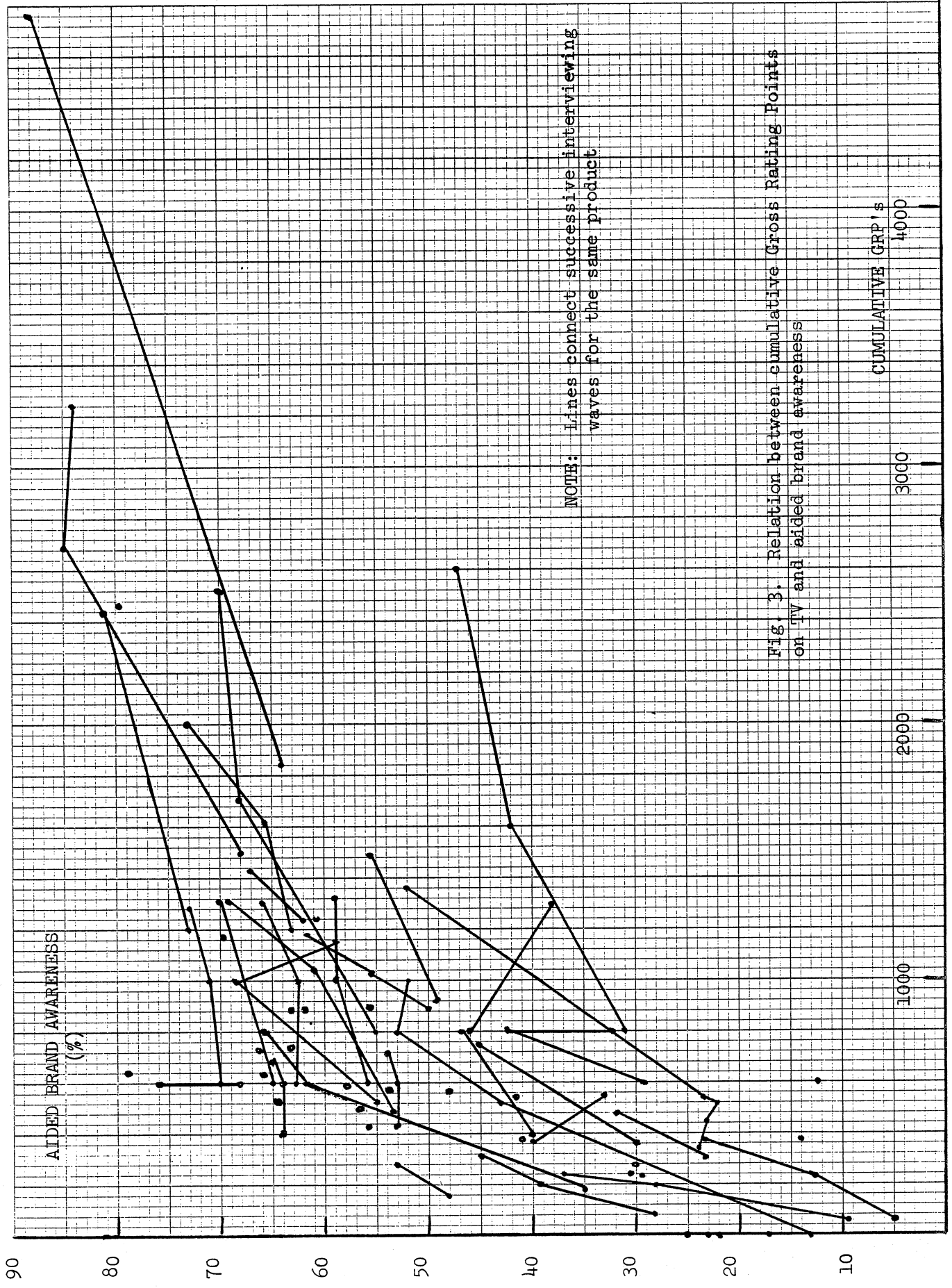
$Y_t$  = Awareness at time  $t$

$S_t$  = Cumulative GRP's at time  $t$

$a$  = Level of awareness attained prior to TV advertising

$b$  = Rate of increase of the awareness function

This curve has a number of interesting features: 1) aided awareness is naturally bounded above by 100%, which products approach very slowly. That is, there are diminishing returns on awareness for increased advertising weight. 2) The Y- intercept is positive, in the range of 10-20%. This



NOTE: Lines connect successive interviewing waves for the same product

Fig. 3. Relation between cumulative Gross Rating Points on TV and aided brand awareness



means that given zero TV GRP's spent behind a new product, one usually obtains 10-20% aided awareness for that product. This is probably due to two main effects: sell-in to stores before the start of advertising, where some consumers may become aware of the brand by seeing it in the store (probably a small effect); and, more importantly, "false" aided awareness for the new product.

Anyone who measures aided awareness of brands is familiar with the concept of "false" awareness. It is assumed that a certain number of respondents who say they are aware of a brand are not "really" aware of the brand in question. Operationally, "false" awareness for a brand can be defined as that level of aided awareness obtained for the product with no advertising, no distribution, and no other marketing effort of any kind.

The whole issue of "false" awareness is the biggest problem with aided awareness, and points to the overall fragility of that measurement of awareness. If the amount of "false" awareness were constant from brand to brand, then there would be no problem, since the curve in Fig. 3 would already incorporate this constant "false" awareness. Data at hand, however, suggest that many brands differ significantly in their levels of "false" awareness.

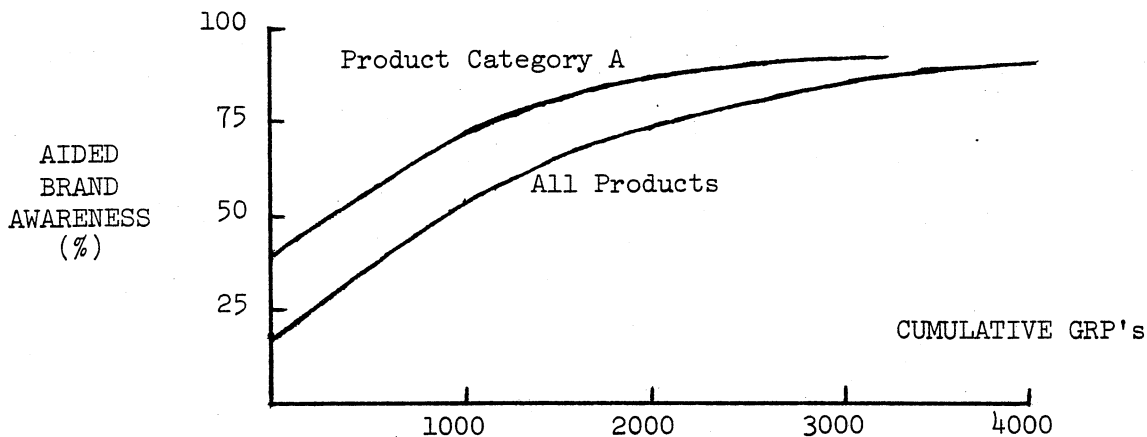


Fig. 4. Relation between cumulative Gross Rating Points on TV and aided brand awareness for all products and for Product Category A.

Figure 4 displays the curve (generated by equation [1]) which best fits the relationship between cumulative GRP's and aided brand awareness, both among all products as a whole, and among products within one particular category. This category is composed solely of new products made by a well-known food producer which are all similar in nature to existing products made by that producer, and which all carried the producer's widely familiar corporate name. Under these conditions, it might well be anticipated that potential exists for substantial "false" awareness of these new products. Figure 4 shows that these specific products do achieve higher levels of aided awareness than most new products (given their spending), and it is reasonable to believe that this difference is primarily due to unusually high levels of "false" awareness for these products. A second indication of the fragility of the measurement of aided awareness is that results can be very dependent upon question order for certain products, especially when "false" awareness exists for those products.

An experiment on question order was conducted while tracking the introduction of a new personal care product. It was known, going into the tracking study, that "false" awareness would probably be a major problem in

interpreting the brand's performance, since its name was very similar to two existing major brands in the category.

The experiment involved rotating the order in which respondents (male product category users) were asked aided awareness of four brands. Two of the brands (A and B) were well-established, high-volume products (with product A the bigger seller), while the other two (C and D) were new products which had similar names to products A and B. It was hypothesized that the levels of aided awareness measured for brands C and D would be higher, due to higher "false" awareness, when they were asked first, before the well-known brands A and B, than when they were asked after brands A and B. It was assumed that consumers who confused the new brands with the old brands would be "helped" by hearing the names of the old brands first.

Table 1 shows the results of this experiment. As expected, the highest levels of aided brand awareness were obtained for the two new products when they were asked first, before the well-known brands were asked, and this difference was significant in magnitude (10-12%). Interestingly, the established products did not show a difference in aided awareness depending on whether they were asked first or last.

TABLE 1

## Aided Awareness, By Order of Asking

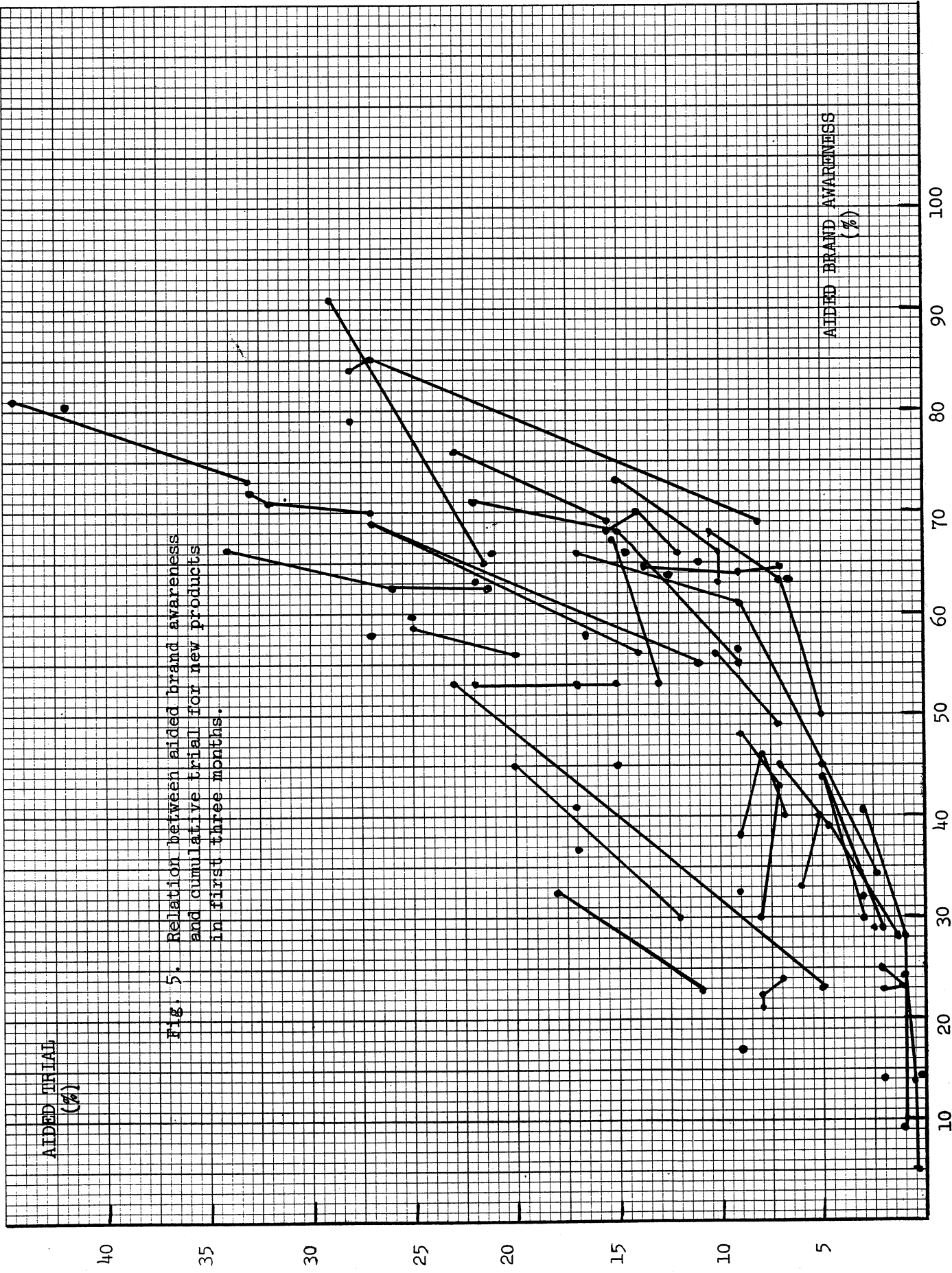
BASE: <u>Adult Male</u> <u>Product Category</u> <u>Users</u>	Brand	Brand	Brand	Brand
	Asked 1st (302)	Asked 2nd (297)	Asked 3rd (301)	Asked 4th (299)
Brand A (established)	93%	95%	94%	92%
Brand B (established)	76	72	73	73*
Brand C (new)	78	77	65*	69*
Brand D (new)	75	63*	63*	61*

\*asked after Brand A (market leader)

This experiment casts great doubt on the usefulness of the measurement of aided awareness, since it appears to be a loose, unstable measurement. Given these problems, how useful is it? The relationship in Figure 3 suggests that it is a useful measurement despite these problems. The effects of these problems do not appear to be so large that general predictive relationships between advertising spending and aided awareness cannot be made. At the same time, it must be understood that aided awareness is not a firm, concrete measurement, and should not be treated as such. The second question to be answered in evaluating aided brand awareness as a relevant measurement is how well does aided brand awareness relate to trial of new products?

Figure 5 shows the relation between aided brand awareness of new products and their trial levels ("normalized" by being computed among potential market for those products) during their first three months after introduction. While again the relationship is by no means perfect, it does appear that a definable relationship exists. Blattberg and Golanty (1974) have quantified this relationship to be

$$T_t = T_{t-1} + \alpha (A_t - A_{t-1}) + \beta (A_{t-1} - T_{t-1})$$



where  $T_t$  is trial in period  $t$ ,  $A_t$  is aided brand awareness in period  $t$ , and  $\alpha$  and  $\beta$  are parameters.

It thus appears that the measure of aided brand awareness is a useful intervening variable between marketing effort (TV GRP's) and trial levels attained by new products. Those relationships, however, are not predictive within a narrow error range. Any specific new product may get much higher awareness levels than the "average" new product (because it has, for example a commercial with more attention-getting power than the "average" new product commercial). But the relationships are useful in predicting potential ranges in performance, given varying levels of advertising weight. When used in this way, aided brand awareness is very definitely a relevant indicator of a new brand's performance, relative to other new products in the past.

#### Misapplications Of Awareness

A whole genre of models designed to predict sales for new products prior to test market involve showing respondents a stimulus (usually a commercial or concept board for the new product) and then measuring, in various ways, trial of the new product. That measure of trial is then factored down for awareness. It is assumed that by showing the stimulus to all respondents, 100% awareness is generated, and that in the real world, 100% awareness will not be obtained for the new product.

While it is clear that any measured trial obtained under these conditions must be factored down to correct for lower real-world awareness, there are some major theoretical problems with the ways this adjustment is typically made.

Most frequently, the adjustment is of the linear form

$$T_A = T \times A \quad (2)$$

where  $T_A$  is the adjusted trial estimate,  $T$  is the measured trial, and  $A$  is some measure of expected awareness (usually aided brand awareness).

There are two basic problems with the use of awareness within this adjustment procedure. First, as has been seen, every measurement of awareness has some significant limitations. Unaided awareness does not seem to be very predictable in terms of advertising weight, and aided awareness is a quite fragile and unstable measurement, subject to problems of "false" awareness and question order. It seems at best unwise to design a model whose prediction of sales varies linearly with variables as uncertain as these.

The second problem with this procedure is philosophical in nature. Which measure of awareness should be used in (2) -- unaided or aided, brand or advertising? Vastly different levels are obtained on each measure for any given product. Which one is proper? In most models of this nature, aided brand awareness is chosen. On what grounds?

Theoretically, the answer is that one must estimate the percent of the population which will be in a state of awareness equal to that state of awareness generated in the artificial, forced-viewing laboratory situation. What might this measurement be?

It could be aided brand awareness. Certainly there is 100% aided brand awareness generated in the laboratory test. But it might also be aided advertising awareness, which is also 100% in such tests. Or perhaps unaided awareness, which might be 100% in such tests. Or even an awareness of "poorer quality" than aided brand awareness, such as recognition of narrative elements in the commercial. The central question is why should one apply aided brand awareness estimates to correct for 100% "awareness"? If awareness is indeed a continuum, as suggested earlier, then any measurement designed to divide the world into those "aware" and those "not aware" is entirely arbitrary. No one measurement is "magical" in capturing the essence of awareness.

If one is going to use "awareness" in these pre-test models, the measurement of awareness used must meet the following criteria:

- 1) It must be estimatable from the marketing plan, within a narrow error range;
- 2) It must have a definable relationship to trial in the real world;
- 3) It must be relatively stable, and not sensitive to variability due to noise (necessary for 1) and 2) to be fulfilled); and
- 4) It must not be highly correlated with those variables which one is trying to ultimately forecast.

The available data on unaided and aided awareness suggest that neither measurement fulfills 1) - 4) above.

The ultimate test of models which hang on this use of the concept of awareness is, of course, whether they work. On theoretical grounds, at least, their potential appears highly doubtful.

#### Summary

Awareness has been shown to be a useful intervening variable between advertising weight and trial for new products. Aided brand awareness is superior to unaided awareness as an intervening variable. At the same time, it must be remembered that awareness is a continuum, and that any measure of it is an artificial dichotomization of reality. Given this fact, as well as the difficulties in actually measuring awareness, the use of awareness to factor down simulated trial measures to correct for real-world awareness is a highly hazardous procedure.

#### FOOTNOTES

1. The author gratefully acknowledges the assistance and support of Dr. Joseph T. Plummer and Mr. Frank Feinberg of Leo Burnett U.S.A. in preparation of this paper.
2. David W. Olson is Associate Research Supervisor, Special Task Force, Leo Burnett U.S.A.
3. The reason for choosing TV advertising weight is that its effects are usually large in comparison to other forms of promotion effort (i.e. most of the marketing dollars are spent in TV). Since the model

described here includes only one input, it is natural to use TV GRP's as that input. For products which are heavily supported by alternative advertising vehicles (e.g., print or consumer promotion), this formulation of the model is inappropriate.

4. Trial throughout this paper will be reported among each product's "potential users." This permits looking jointly at products in many different categories, since their trial levels are "normalized" by basing them on their own potential market.
5. This, of course, leads to another methodological problem with unaided awareness, which is: what does one do with a new product which does not fit cleanly into an existing product category?
6. This last point is not strictly true, of course. Some people can and do become aware of a new product (as measured by aided awareness) only after trying the product, which violates the causal structure of the operational model. The magnitude of this effect, however, is usually small enough to be negligible.

#### REFERENCES

- Assael, H. & Day, G.S. Attitudes and awareness as predictors of market share. Journal of Advertising Research, 1968, 8 (4), 3-10.
- Axelrod, J.N. Attitude measures that predict purchase. Journal of Advertising Research, 1968, 8 (1), 3-17.
- Blattberg, R. & Golanty, J. An early test market forecasting model for new product planning. Working paper, 1974.
- Gruber, A. Top-of-mind awareness and share of families: An observation. Journal of Marketing Research, 1969, 4, 227-231.
- Palda, K.S. The measurement of cumulative advertising effects. Englewood Cliffs, N.J.: Prentice-Hall, 1964.



EFFECTS OF POLARITY OF SEMANTIC DIFFERENTIAL SCALES  
IN CONSUMER RESEARCH

John Dickson  
University of Connecticut

Gerald Albaum  
University of Oregon

This paper examines the results of a laboratory type experiment designed to study the question of the effects of polarity of semantic differential scales. Polarity refers to the horizontal position of the positive (i.e., favorable) and negative adjective or phrase used for each bipolar scale in a semantic differential test. A total of 84 subjects rated two concepts across a set of 20 bipolar scales, which were phrases. Each subject randomly received one of three treatment variations of a semantic differential measurement instrument.

This study provides evidence which indicates that the polarity of a scale does not significantly affect the results. Since it makes little difference how groups of scales are presented, by placing all positive or negative phrases on the same side, coding, editing, and processing errors and costs should be reduced.

#### Introduction

The use of various opinion and attitude measuring devices by marketing researchers has become increasingly popular in recent years. Of particular interest to researchers are those devices which permit a comparison of attitudes about product images, advertisements and other marketing concepts. One of the more widely used devices for collecting such information is the semantic differential, developed by Osgood and his associates (1957). Other scaling techniques such as multidimensional scaling, the Likert scale, and the Stapel scale (Hawkins, 1973) have also been used for similar purposes.

Yet, the increasing popularity of utilizing the various scaling models for measuring attitudes and opinions has led several researchers to explore more thoroughly certain methodological questions. Some of the current issues being raised include the appropriate number of points to use in presenting a scale (Green and Rao, 1970; Jacoby and Matell, 1971; Lehman and Hulbert, 1972), inclusion of "ideal" ratings (Landon, 1971), the question of forced-choice (Albaum and Munsinger, 1973; Hughes, 1969), concept-scale interaction (Sharpe and Anderson, 1972), the way in which scales are administered (Albaum and Munsinger, 1973; Hawkins, 1973), reversing the order of presentation of scales (Belson, 1966; Payne, 1972), and polarity of statements (Falthzik and Jolson, 1974). Although the published research concerning the above questions covers a variety of types of scaling techniques, all issues are directly applicable to the semantic differential.

The purpose of this paper is to present the findings of a study concerned with one specific methodological question--the polarity of a semantic differential scale. By polarity we mean the horizontal position of the positive



(i.e., favorable) and negative adjective or phrase used for each bipolar scale in a semantic differential test.

Considering reading patterns of people in the United States and similar cultures, for any given set of semantic differential bipolar scales three alternative arrangements concerning polarity are possible:

1. Positive: all scales have the "positive" adjective or phrase on the left-hand side of the scale.
2. Negative: all scales have the "negative" adjective or phrase on the left-hand side of the scale.
3. Mixed: some of the scales have the "positive" adjective or phrase on the left-hand side while others have the "negative" adjective or phrase on the left-hand side.

Since it is believed that the formation of position preferences may create unwanted biases, most researchers imply or state that scales representing the same underlying factor should be randomly alternated in direction of polarity (Tull and Albaum, 1973; Osgood et al., 1957). Typically, published research describes which and how many scales were reversed in polarity, thus implicitly agreeing with the above assumption. However, a thorough search of the literature yielded no information or empirical research which substantiates this assumption. Consequently, a laboratory-type experiment was conducted to test the following hypotheses:

Hypothesis I: There is no significant difference in the results obtained from a semantic differential measurement instrument in which the polarity of the scales is "positive" and one in which the polarity of scales is mixed.

Hypothesis II: There is no significant difference in the results obtained from a semantic differential in which polarity is "negative" and one in which polarity is mixed.

Hypothesis III: "Positive" and "negative" semantic differential instruments yield the results which are not significantly different.

#### Methodology

The treatment variable was a semantic differential measurement instrument consisting of twenty seven-point bipolar scales. Subjects were asked to rate two concepts, which were department stores. If the concepts yield similar data, increased confidence in the findings would result. In addition, two concepts of the same general type can serve as a test for the existence of an unusual concept-subject interaction.

Nine of the scales used were selected from among those previously tested and recommended by Kelly and Stephenson (1967), while eleven scales were selected from an earlier unpublished study conducted by one of the authors in which similar type concepts were analyzed. Phrases, rather than adjectives, were used to increase the relevance of the scales to the retailing context. This modification of the usual semantic differential model was suggested by Mindak (1961) for situations where concepts may provide an environment in which concept-scale interaction may occur due to the lack of relevance to subjects of the adjectives developed by Osgood et al. (1957).

Three treatment measurement instruments were administered:

1. mixed arrangement of polarity in which one-half the scales had the positive phrase, and one-half the negative, on the left side. Assignment of scales to these two groups was done on a random basis.
2. positive arrangement of scale polarity.
3. negative arrangement of scale polarity.

The vertical ordering of each scale used in the treatments was assigned randomly and was the same for all treatments. Also, for both concepts, the ordering of the bipolar scales was held constant. The order of presentation of the concepts was such that one-half the subjects rated Store A first while the other subjects rated Store B first.

Upper division undergraduate business students attending the University of Oregon were used as subjects. The concepts used were selected to insure that all of the subjects would be familiar with them. This was verified during the test. Since the function of this sample of students was to provide answers to a methodological question, there is no reason to believe such a sample would create bias problems. The treatments were assigned randomly to students in two classes. The results for each class were compared for significant differences in response distributions using the chi-square test. In addition, the Mann-Whitney U test was used to test for significant differences in location (i.e., central tendency). No significant difference was observed in terms of distribution or direction of responses so the two groups were combined for data analysis. The combined sample consisted of 84 subjects, distributed among the treatments as follows:

<u>Treatment</u>	<u>Number of Subjects</u>
1 (mixed)	28
2 (positive)	27
3 (negative)	29

Precautions were taken to ensure that students enrolled in both classes participated only once.

### Results and Discussion

Concepts were analyzed separately in two different ways. First, each scale was treated independent of the others and treatment comparisons made as follows: (1) mixed vs. positive polarity, (2) mixed vs. negative, (3) positive vs. negative. The Mann-Whitney U test was used for the statistical analysis of these paired data. In some situations, researchers may be interested in deriving summated scale scores for each individual and then comparing the scores for one group of individuals to a second group (Osgood, *et al.*, 1957). A second form of analysis was conducted to determine whether significant differences exist among the summated scores of individuals subjected to the three treatments. Each subject was assigned a score for each concept according to the sum of his responses on all scales. To test if there was a significant difference among treatments, the Kruskal-Wallis one-way analysis of variance test was applied to the mean summated scores for the groups of subjects exposed to the three treatments.

### Individual Scale Analysis

The results of the paired-treatment comparisons for each scale are shown in Tables 1 and 2, which also present treatment mean values (scales were scored 1 to 7, positive to negative). The significance test results are summarized in Table 3, which indicates that in all situations except the positive-negative comparison for Store A the number of significant scales ( $\alpha = .10$ ) are less than the maximum number that would be expected to be significant by chance alone, on the basis of the binomial distribution. For example, the probability of finding 3 significant scales at  $\alpha = .10$  out of 20 possible by chance alone is .32. These data indicate that Hypotheses I and II cannot be rejected.

The results relevant to Hypothesis III are mixed. That is, for Store A the number of significant scales ( $\alpha = .10$ ) for the positive-negative treatment comparison exceeds chance expectations whereas for Store B the number can be attributed to chance. Consequently this analysis cannot be used for reaching a definite conclusion concerning Hypothesis III.

### Summated Score Analysis

As indicated above, the responses for each subject were summed over all 20 scales. The mean values of summated scores for each treatment are presented in Table 4. These data were analyzed by the Kruskal-Wallis one-way analysis of variance to test for significant differences in the summed values. As shown, the probability at which significance occurs exceeds .10 for both concepts. Consequently, a test of the three hypotheses with  $\alpha = .10$  as the critical value would show that all three cannot be rejected.

### Conclusions

The data in this study indicate that there are no statistically significant differences in results between a semantic differential in which all positive and all negative phrases are placed on one side and the more typical situation where each side has one-half positive and one-half negative. Both the individual scale and summated score analyses support this contention. Unfortunately, the results are inconclusive for indicating which is better, placing all the positive phrases on the left side or all the negative phrases there.

Our findings do suggest some rather pragmatic values to researchers. Since this study showed that it makes little difference how a group of scales are presented, by placing all positive and all negative phrases on the same side in a semantic differential, coding, editing, and processing errors and costs should be reduced. In addition, preparation of the instrument in the first place may be facilitated. Moreover, since many adjectives and phrases are difficult to classify as positive or negative, if polarity is not an issue then there is one less matter for the researcher to be concerned about.

As in most exploratory-type research projects, various limitations to our findings are imposed as a result of methodological problems as well as the inherent nature of the test design. First and foremost, since students were used as respondents, the findings cannot really be extrapolated beyond the student population. In essence, the convenience sample of students (randomly assigned to treatments, however) results in the test design being a fixed effects model.

TABLE 1  
Treatment Mean Values and Paired-Treatment Comparisons for Store A, by Individual Scale

Scale (Positive Phrase Only)	Treatment <sup>a</sup> Mean Values		Negative (N=29)	Significance of Comparisons (Mann-Whitney U)	
	Mixed (N=28)	Positive (N=27)		Mixed vs. Positive	Mixed vs. Negative
1. * Fast check out	4.00	3.35	4.06	p<.10	p<.05
2. Adequate parking	4.48	4.07	4.31		
3. Helpful personnel	3.59	3.08	3.79		p<.10
4. *Store emphasizes service	3.74	3.61	4.00		
5. *Wide selection of products	3.93	3.19	3.58	p<.10	
6. *Attractive store appearance	2.81	2.73	3.10		
7. *High quality products	2.96	2.80	2.79		
8. Friendly personnel	3.48	3.23	3.76		
9. *Informative advertising	4.14	3.84	4.17		
10. Always improving	3.92	4.07	4.24		
11. Convenient location	3.85	3.80	3.72		
12. *Well known	2.56	2.50	3.13		
13. High values for money spent	4.37	3.35	3.93	p<.05	p<.10
14. Fully stocked	3.86	3.15	3.85		p<.10
15. *Easy to find items you want	3.48	3.54	4.14		
16. *Well liked by your friends	3.63	3.50	3.90		
17. Large number of specials	4.41	4.12	4.34		
18. *Advertising frequently seen by you	4.63	4.12	5.24		p<.05
19. Believable advertising	3.63	4.00	3.76		
20. Low prices compared to other stores	5.00	5.19	5.03		

\* These scales had the positive phrase on the left side of the mixed polarity treatment.

<sup>a</sup> Scales were scored 1 to 7, positive to negative.

TABLE 2

## Treatment Mean Values and Paired-Treatment Comparisons for Store B, by Individual Scale

Scale (Positive Phrase Only)	Treatment Mean Values <sup>a</sup>			Significance of Comparisons (Mann-Whitney U)	
	Mixed (N=28)	Positive (N=27)	Negative (N=29)	Mixed vs. Positive	Mixed vs. Positive vs. Negative
1. *Fast check out	3.54	3.19	3.55		
2. Adequate parking	1.86	1.41	2.10		
3. Helpful personnel	3.32	3.03	3.51		
4. *Store emphasizes service	4.04	3.41	3.97	p<.10	
5. *Wide selection of products	2.68	2.76	2.75		
6. *Attractive store appearance	2.75	2.96	2.52		
7. *High quality products	4.00	3.30	3.72	p<.05	
8. Friendly personnel	3.03	3.19	3.48		
9. *Informative advertising	3.64	3.69	3.55		
10. Always improving	3.82	3.52	3.76		
11. Convenient location	2.96	2.37	2.96		
12. *Well known	1.86	1.61	1.72		
13. High values for money spent	3.36	3.38	3.24		
14. Fully stocked	2.96	2.67	3.10		
15. *Easy to find items you want	3.71	3.35	3.55		
16. *Well liked by your friends	4.00	3.59	3.45		
17. Large number of specials	3.78	3.46	3.93		
18. *Advertisements frequently seen by you	3.67	3.31	3.48		
19. Believable advertising	3.39	3.49	3.38		
20. Low prices compared to other stores	3.39	3.12	3.20		p<.10

\* These scales had the positive phrase on the left side of the mixed polarity treatment.

<sup>a</sup> Scales are scored 1 to 7, positive to negative.

TABLE 3

Number of Significant Scales at  $\alpha = .10$ , by Concept

Treatment Pairs	Store A	Store B
Mixed-Positive	3	2
Mixed-Negative	1	0
Positive-Negative	6 <sup>a</sup>	1

<sup>a</sup>The probability of this many significant differences ( $N=20$ ,  $\alpha = .10$ ) occurring by chance is less than .02 based on the binomial distribution.

TABLE 4

Mean Values of Summated Scores

Concept	Treatment			Chi-Square
	Mixed	Positive	Negative	
Store A	76.47	71.27	78.82	3.71 ( $p>.10$ )
Store B	65.76	60.81	65.10	3.94 ( $p>.10$ )

Secondly, no pretest was conducted using the specific questionnaire which was used in this test. Most of the bipolar phrases were selected on the basis of recommendations by other authors. However, one of the present authors has used all the scales in a previous test, although the scales were arranged in a different order.

Finally, as in the case of all semantic differentials where there are more than one scale per concept, there may be a lack of independence between scales, and there may exist a concept-scale interaction effect. However, we believe such confounding effects did not appear in this study.

## REFERENCES

- Albaum, G. & Munsinger, G. Methodological questions concerning the use of the semantic differential. Unpublished paper presented at the Southwestern Social Science Association meeting, March, 1973.
- Belson, W. A. The effects of reversing the presentation order of verbal rating scales. Journal of Advertising Research, 1966, 6, 30-37.

- Falthezik, A. M. & Jolson, M. A. Statement polarity in attitude studies. Journal of Marketing Research, 1974, 11, 102-105.
- Green, P. E. & Rao, V. R. How many rating scales and response categories to use. Journal of Marketing, 1970, 34, 33-39.
- Hawkins, D. I. Utilization of the Stapel Scale in attitude measurement. Unpublished paper presented at the Southwestern Social Science Association meeting, March, 1973.
- Hughes, G. D. Some confounding effects of forced choice scales. Journal of Marketing Research, 1969, 6, 223-226.
- Jacoby, J. & Matell, M. S. Three-Point Likert Scales are good enough. Journal of Marketing Research, 1971, 8, 495-500.
- Kelly, R. F. & Stephenson, R. The semantic differential: an information source for designing retail patronage appeals. Journal of Marketing, 1967, 31, 43-47.
- Landon, E. L. Order bias, the ideal rating, and the semantic differential. Journal of Marketing Research, 1971, 8, 375-378.
- Lehmann, D. R. & Hulbert, J. Are three-point scales always good enough? Journal of Marketing Research, 1972, 9, 444-446.
- Mindak, W. A. Fitting the semantic differential to the marketing problem. Journal of Marketing, 1961, 25, 28-33.
- Osgood, C. E., Suci, G. J. & Tannebaum, P. H. The measurement of meaning. Urbana, Ill.: University of Illinois Press, 1957.
- Payne, J. D. The effects of reversing the order of verbal rating scales in a postal survey. Journal of the Marketing Research Society, 1972, 14, 30-44.
- Seigel, S. Nonparametric statistics. New York: McGraw-Hill, 1956.
- Sharpe, L. I. & Anderson, W. Concept-scale interaction in the semantic differential. Journal of Marketing Research, 1972, 9, 432-434.
- Tull, D. S. & Albaum, G. Survey research: A decisional approach. New York: Intext Educational Publishers, 1973.

## BOOTSTRAPPING OF DATA AND DECISIONS

Joel Huber  
Purdue University

Bootstrapping involves the substitution of a simple linear model of judgments in place of the judgments themselves. It has been found that in many decision-making contexts the bootstrapped decisions are better than the judgments from which they were derived. It appears that the linear model is quite successful at capturing the policy of the judge and then making decisions without human inconsistency. Most of the work done on bootstrapping has been done in a context--such as forecasts--where the criterion of accuracy is clearly defined. This study shows that bootstrapping can be used to upgrade the quality of subjective judgments (data) which have no ultimate criterion of accuracy but are judged in terms of their usefulness as input data to a predictive model. Implications are explored as to the use of bootstrapping of both data and decisions in consumer behavior.

Models of judgment which associate overall worth with a linear combination of the affective and cognitive components are certainly familiar to those involved in consumer research. For example, these models are often used as diagnostic tools by manufacturers interested in changing the image or physical make-up of a product. Bootstrapping represents a different posture toward the modeling of judgments. Linear models are used to replace the raw judgments, rather than, in any rigorous sense, to understand them. Judgments or decisions are represented as linear combinations of cues where the weights are typically derived from a multiple regression using the judgments as the criterion and the cues as predictors. The ensuing decompositional model provides a linear approximation to what the subject is doing as can be inferred from judgments and inputs. Hoffman (1960), who was the first to formally suggest such a procedure, termed such a linear approximation a "paramorphic representation" of the judgment process. These models are paramorphic in the sense that while they might predict quite well, they are not to be construed as models of what the judge is really doing. That is, decision makers may behave as if they are linear machines but that does not mean that, in fact, they are.

When judgments are made in a context where they can be compared with the "true" value, as in a forecast, the linear model has generally been found to be more accurate than the raw judgments. That is, the correlation of the ultimate criterion with the linear model is generally greater than its correlation with the raw judgments used to derive the model. The replacement of raw judgments with this linear combination of cues has come to be called "bootstrapping," a term coined by R. M. Dawes. In effect, the raw judgments "lift themselves up by their bootstraps." Two examples should be sufficient to illustrate this technique.



One of the first studies to suggest that simple linear models of decisions might produce good decisions was Yntema and Torgerson (1961). Subjects were provided with ellipses of different size, shape and color. Worth was defined so as to increase nonlinearly with increases in size, thinness and brownness. After a ten day training period consisting of giving subjects feedback on their predictions of worth, subjects were required to make a battery of 180 judgments without feedback. The average product moment correlation between these judgments and the true worths was 0.84. A simple additive bootstrapping model derived from predicting these judgments as functions of size, shape and color yielded an average correlation with the true values of 0.89. Thus, the additive bootstrapping models were more accurate than the judges in spite of the fact that these models could not take into account interactions, while the human judges presumably could.

Goldberg (1970) used judgments of clinical psychologists to build a model to discriminate neurotics from psychotics. The predictor variables were scores on the Minnesota Multiphasic Personality Inventory, a test which provides a profile of patients along 11 dimensions. The bootstrapping model provided superior predictions of later diagnosis for 26 out of the 29 judges. Similarly, Dawes (1971) found that admissions evaluations have a higher correlation with actual achievement if derived from a bootstrapping model of judgments rather than the judgments themselves. Bowman (1963) and Kunreuther (1969) were able to demonstrate improved decisions in the field of production management, while in marketing, Heeler *et al.* (1973) and Montgomery (1972) have applied bootstrapping to the decisions of buyers for supermarket chains with similar results.

In conclusion, then, different researchers, working in different fields, have found that a simple linear model of judgments satisfies the objectives of the judge better than the original judgments.

#### Why Bootstrapping Works

Bootstrapping works because the linear model is able to make extremely good approximations of most decision processes. The model then makes these judgments without random error. Thus by bootstrapping one replaces the random error of the judge with the nonrandom error of the model. This nonrandom error can be broken into two components: (1) a calibration error due to insufficient sample size to reliably estimate the parameters and (2) a specification error due to the inability of the linear model to capture the complexities of what the judge is doing. These two sources of error will be discussed in an attempt to explain why their sum has generally been less than the error of raw judgments.

Calibration error reflects a fairly minor component of the error in a bootstrapping model. Dudycha and Naylor (1966) provided subjects with cues and worths that were related by linear model with different levels of error. After a learning period, the bootstrapped models of judgments on 50 stimuli had average correlations with the optimal model of better than 0.90. Even where the beta coefficients appear to be unstable themselves due to multicollinearity, the predictions from such a model tend to be quite stable.

Specification error reflects the inability of the linear model to account for nonlinearities or interactions in the judgment process. The first researchers of paramorphic representation (concentrated mainly at Oregon Research Institute) saw the linear model as a first approximation of the

decision process which would be modified later by nonlinear and interactive components. Then a funny thing happened. Adjustments to the linear model provided very little improvement to predictive accuracy. This result was anticipated by Yntema and Torgerson (1961) who found that a main effects model  $Y = a_i + b_j + c_k$  accounts for over 90% of the variance of data generated by the multiplicative model  $Y = ij + ik + jk$  (where  $i$ ,  $j$ , and  $k$  are integers between one and seven). Thus, one loses very little by approximating such a decision process with only the main effects.

In a large simulation study Rorer (1971) tested the ability of the linear model to approximate data generated by interactive and configural models. These included interactive and configural terms as well as disjunctive and conjunctive step functions and elaborate lexicographic models. The linear model was generally able to account for over 80% of the variance. Furthermore, given reasonable levels of error, the interaction terms in most cases would not be significant. This result appears to be quite general as long as the criterion variable is conditionally monotone with respect to the cues. That is, if the direction of the effect of a cue is the same regardless of the levels of the other cues. Cues which are conditionally monotone appear quite often in judgmental situations. For example, economy, performance, styling and closeness to mid-sized are all attributes which in a rational man might be conditionally monotone with his judgments of overall worth of automobiles. Research has shown that if this is the case then a linear model will do a good job of approximating the judgments. Furthermore, with fallible data the analysis of variance will generally lack the power to measure the incremental gain of a non-linear analysis.

While the bootstrapping of decisions generally produces better results than the decisions from which it was derived, it generally does not produce the best linear decision scheme available. Simply regressing the cues directly on the criterion produces better results than going through judgments (Meehl, 1954). In fact, Dawes and Corrigan (1974) show that linear models with random coefficients (but the correct sign) do as well as the bootstrapping models. Thus bootstrapping models are not to be seen as magical or in any sense optimal linear models, but merely a method for picking the appropriate variables and weighing them in the right direction. Furthermore, if an unambiguous criterion exists, a better model can be derived by regressing the cues directly on it.

It could be argued that there are many situations where the "optimal" linear model derived from a regression of cues on the criterion is less valid than the bootstrapping model derived from a regression of cues on decisions. Consider the admissions problem. Bootstrapping does not do as well as an optimal linear model in predicting rank in class. However, it is feasible that the admissions committee is taking into consideration other goals, such as racial balance or being well-rounded. These considerations would be reflected in the coefficients of the bootstrapping model but not in a model calibrated to class rank. Thus if the objective is to provide a model that satisfies the judge's implicit goals, bootstrapping provides at least a first step in this direction. It is this quality that makes bootstrapping particularly appropriate to many problems in consumer behavior.

### Data Bootstrapping

The typical validations of bootstrapping have used judgments, such as forecasts, for which the ultimate criterion for accuracy is easily specified. Further, the cues that go into the judgments have been clearly specified and are generally of quantitative form. This study considers the applicability of bootstrapping to data that serves as input to behavioral models and generally lacks the above qualities. A response of a subject to a stimulus or question cannot have ultimate validity but only be considered better or worse to the extent that it can be related to other responses or behavior on the part of the subject. For example, the superiority of a measure of intention to purchase can be ascertained on the basis of its correlation with actual purchases. In the same way bootstrapping will be evaluated on the basis of its effectiveness in improving input to a behavioral model.

The present study represents an attempt to evaluate bootstrapping on preference judgments of particular samples of iced tea. All analysis is done on the basis of the individual subject. The test between bootstrapped and raw judgments is made by comparing which provides better predictions of preference.

### Experimental Design

The preference judgments from a convenience sample of 22 people were used for this study. Each was required to make judgments on samples of Lipton iced tea that differed over the amount of sugar and tea according to a balanced design. As is illustrated in Figure 1, they were required to make judgments on 7 validation stimuli nested within 16 calibration stimuli.

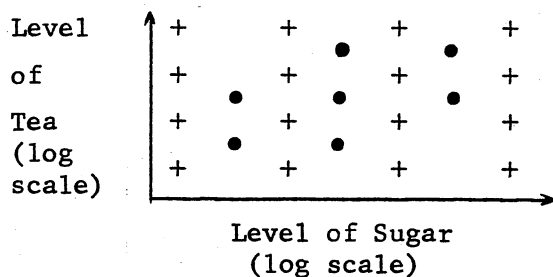


Figure 1. Relationship of 7 validation stimuli (•) embedded in 16 calibration stimuli (+)

For each subject the analysis revolved about the following data.

$P_i$  = Preference scale for stimulus  $i$ ,  $i=1,16$  for the calibration stimuli and  $i=17,23$  for the validation stimuli. This scale was formed for each set from preference differences using Scheffé's (1951) method of analysis modified for analysis of individual data.

$\delta_{ik}$  = Judgment as to the degree to which stimulus  $i$  has too much, or too little, sugar ( $k=1$ ) or tea ( $k=2$ ). These were coded on an integer scale from -3 to +3, negative numbers indicating too little, zero indicating optimum, and positive numbers indicating too much of the ingredient.

$x_{ik}$  = Objective level of sugar ( $k=1$ ) and tea ( $k=2$ ) for stimulus  $i$ .

Using data on the calibration stimuli, a preference function

$$(1) \quad P_i = f(\delta_{ik}) = b_0 + b_1 |\delta_{i1}| + b_2 |\delta_{i2}|$$

is estimated by multiple regression for each individual. This is a version of the familiar weighted-additive model. The absolute value of the  $\delta_{ik}$ 's can be interpreted as the distance from stimulus  $i$  from the ideal along dimension  $k$ . Thus preference is assumed to be a function of the sum of weighted distances along these psychological dimensions. This parameterized model is then used to make predictions on the preferences of the seven validation stimuli.

The bootstrapping model relates the  $\delta_{ik}$ 's to the real levels of sugar and tea. This is for dimension  $k$

$$(2) \quad \delta_{ik} = g(x_{ik}) = b_{0k} + b_{1k}x_{i1} + b_{2k}x_{i2}$$

The effectiveness of bootstrapping is gauged by whether estimates ( $\hat{\delta}_{ik}$ 's) generated from Equation 2 produce better predictions as input to Equation 1 than the raw data ( $\delta_{ik}$ 's).

The bootstrapping model given in Equation 2 assumes the amount of change desired in tea and sugar is a linear function of the actual values of these variables. If, for a given individual, preferences are single-peaked or monotone within the physical space, then the physical levels will be monotone with the  $\delta_{ik}$ 's. Furthermore, since linear functions have been shown (Rorer, 1971; Dawes and Corrigan, 1974) to produce close approximation to most monotone functions, the linear model appears reasonable in this case. This is further supported by the fit of the calibration stimuli to Equation 2. The average product moment correlation across subjects was .84 for sweetness and .62 for tea.

#### A Test of Data Bootstrapping

The effectiveness of bootstrapped against raw judgments was compared at two junctures of the prediction process: (1) to parameterize the preference equation and (2) as input to the parameterized models. In both cases the predicted  $\delta_{ik}$ 's were simply substituted for the raw  $\delta_{ik}$ 's to test bootstrapping. As is shown in Table 1, using bootstrapping to parameterize has relatively little effect while its use on models that have been parameterized produces large and significant gains in prediction.

If one considers the bootstrapping equations to be the first stage in a two-stage model, then using bootstrapping to parameterize the preference model is equivalent to two-stage least squares. This procedure has some theoretical advantages in that errors of the bootstrapped values are not correlated with the error terms of the preference scores. In this case, however, the two-stage model did not produce significant gains probably because the errors in the  $\delta_{ik}$ 's are relatively random and because of the well-known robustness of linear regression to random error in the independent variables.

By contrast, using bootstrapping to produce variables as input to the parameterized models produced large gains in predicting the preferences on

TABLE 1

Effect on Prediction of Using Bootstrapping Measures  
of Subjective Sugar and Tea for Parameterizing  
and as Input to Predictive Model

		Bootstrapping used to Parameterize		
		NO		YES
Bootstrapping Used in Model for Prediction	NO	.70	Significance of Difference *	.70
	YES	.77		.80

Horizontal arrows between columns: NO to YES (top) is NS; YES to YES (bottom) is NS.  
 Vertical arrows: NO to YES (left) is  $p < .10$ ; YES to YES (right) is  $p < .05$ .  
 Diagonal arrows: NO to YES (top-right) is  $p < .05$ ; YES to NO (bottom-left) is  $p < .10$ .

Root-mean-square of product moment correlation  
of predicted to actual preference scale.

\*The significance of the difference in the predictive fit of a pair of models is done by taking the difference between the Fisher (1925) z-transform of the correlations for each subject. A t-test is then used to test the null hypothesis that this difference is uniformly zero across the 22 subjects in the study.

the validation stimuli. This result could only have occurred if the bootstrapped values were, in fact, more accurate estimates of subjective sugar and tea than the original data. This technique could be used to upgrade the quality of data in a wide variety of behavioral science contexts. For example, judgments of sportiness in automobiles could be bootstrapped as a linear function of speed, acceleration, width-to-height ratio, and cornering ability. As input to a second model involving product choice, the bootstrapping equations would provide not only greater reliability but also assistance to product designers interested in translating "sportiness" into more objective components.

#### Discussion

The use of a model to impart reliability to raw data is hardly novel in behavioral research. The technique perhaps most closely related to the above use of data bootstrapping is spatial or temporal smoothing. Instead of assuming a global linear relationship, these techniques assume local linearity so that each point can be approximated by a simple function of several contingent points. MacKay (1973) used spatial smoothing to impart greater

reliability to store-usage data and found that this resulted in better fits and more interpretable solutions to the quadratic regressions which were used to produce market penetration maps. Even in cases where there is not an objective physical or temporal dimensionality, the assumption of linearity is used to improve data quality. Factor analysis assumes a linear relationship between variables and uses this to impart redundancy in the factors. In the same way the bootstrapping equations used in this study can be seen as the use of a particular linearity assumption that increases the reliability of any particular judgment by forcing it to be consistent with the other judgments in the set.

There is a simple preliminary test to determine whether a bootstrapping model will upgrade the quality of data. It requires one replication of the original judgments. A bootstrapping model calibrated on each half is validated against the other half. The average correlation is called the "double cross-validated correlation." This is compared to the reliability of the judgments which is simply the correlation between the two halves. If the model is exact, the cross-validated correlation should approach the square-root of the reliability. It will not be exactly so unless the sample of judgments is infinite. For practical purposes, however, if the cross-validated correlation is greater than the reliability of the judgments, then the bootstrapping model is better at predicting the raw data itself. There are some risks attached to this procedure. One is that the two replications may not be independent draws from the same judgment process but may reflect a change of viewpoint or merely a rote repetition of the first judgments. The second problem is that the bootstrapped relationship, however optimal it might be with respect to the judgments, may not be at all optimal with respect to the criterion or use to which it is being put. In Einhorn's (1972) study, the optimal combination of cues to predict the criterion was disjunctive while the optimal bootstrapping model was conjunctive.

To avoid just these kinds of problems Kunreuther (1969) advocates the use of bootstrapping only when (1) the decision rules are constant over time, (2) the model and resultant coefficients make theoretical sense and, (3) are statistically significant. While these rules apply more in a production programming context than the present one, certainly the call for a priori consideration of the bootstrapping model is valid.

#### Implications for Consumer Behavior

The decomposition models that have been considered here as bootstrapping models are likely to be as good at "capturing" consumer decisions as they have been of capturing the decisions of admissions committees, psychologists and production planners. It is also likely that the same frustrating search for interactions and mediator variables that has occurred in these areas will be paralleled in research on consumer behavior.

The user of bootstrapping, to the extent that he is more concerned with approximating than understanding decision processes, is content to use the robust linear models and rest secure in the knowledge that rare combinations of cues and nonlinearities have little effect on explained variance. Such predictive, decompositional models can, however, be quite useful in the study of consumer behavior. They could be used as a preliminary normative step to enable the consumer to understand the implications of his own decision processes. Alternatively, they could be used by regulatory agencies as a first approximation of what the consumer decision process is. Values then might be

inferred from decisions rather than imposed from above.

Data bootstrapping is likely to have more limited use in consumer behavior. It can be seen as one of a number of methods to provide reliability in data through the structure of a model. Furthermore, just as such models should be consistent with existing theory, they can add to it by explaining, at least in a preliminary way, the sources of the raw judgments. Thus, data bootstrapping can provide more than just reliable data, it can provide a basis for a understanding of its own validity.

#### REFERENCES

- Bowman, E. H. Consistency and optimality in management decision making. Management Science, 1963, 9, 310-321.
- Dawes, Robyn M. A case study of graduate admissions: Application of three principals of human decision making. American Psychologist, 1971, 26, 180-188.
- Dawes, Robyn M. & Bernard Coorigan. Linear models in decision making. Psychological Bulletin, 1974, 18, 95-106.
- Dudycha, L. W. & J. C. Naylor. Characteristics of the human inference process in complex behavior situations. Organizational Behavior and Human Performance, 1966, 1, 110-128.
- Fisher, R. A. Statistical methods for research workers. Edinburgh: Oliver and Boyd, 1925.
- Goldberg, L. R. Man versus model of man: A rationale, plus some evidence, for a method of improving on clinical inferences. Psychological Bulletin, 1970, 73, 422-432.
- Heeler, Roger M, Michael J. Kearney & Bruce McHaffey. Modeling supermarket product selection. Journal of Marketing Research, 1973, 10, 34-37.
- Hoofman, P. J. The paramorphic representation of clinical judgment. Psychological Bulletin, 1960, 57, 116-131.
- Kunreuther, Howard. Extentions of Bowman's theory of managerial decision making. Management Science, 1969, 15, 8.
- MacKay, David B. Spatial measurement of retail store demand. Journal of Marketing Research, 1973, 10, 447-453.
- Meehl, P. E. Clinical versus statistical prediction: A theoretical analysis and review of the literature. Minneapolis: University of Minnesota Press, 1954.
- Montgomery, David B. New product distribution: An analysis of supermarket buyer decisions. Marketing Science Institute Research Programs, Cambridge, Mass., 1973, 63.

- Rorer, L. G. A circuitous route to bootstrapping. In H. B. Haley, A. G. D'Costa and A. M. Schafter (Eds.), Conference on Personality Measurement in Medical Education. Wash., D.C.: Association of American Medical Colleges, 1971.
- Scheffe, H. An analysis of variance for paired comparisons. Journal of the American Statistical Association, 1952, 47, 381-400.
- Yntema, D. B. & W. S. Torgerson. Man-computer cooperation in decisions requiring common sense. IRE Transactions of the Professional Group on Human Factors in Electronics, 1961, HFE-2(1), 20-26.





SERVICE ANALYSIS: A BANK MARKETING EXAMPLE  
USING PERCEPTUAL MAPPING<sup>1</sup>

William Gillette<sup>2</sup> and Richard H. Evans<sup>3</sup>  
Syracuse University

This study involved a form of perceptual mapping and its application to bank marketing. Data was collected on bank service saliency and perceptions. The factor analysis technique was used to position respondents and banks on an n-dimensional map. The results of the study were such that in all but one of the perceptual maps the ideal point (service salience) of the respondents was in a different quadrant than the perceptions of the banks. In addition, the perceptual maps indicated that the banks form a cluster in varying degrees regarding their perceived services. The paper discusses how service re-positioning, market segmentation, and service surveillance may be considered in light of the above findings.

Perceptual mapping is a generic term that applies to the positioning of perceptual points in multidimensional space. The technique has been used in various ways in the past. As an example, Green and Carmone (1969) analyzed automobiles and consumer opinion, Assael (1971) focused on advertising, Lehmann (1972) delved into brand switching and Johnson (1971) indicated how perceptual mapping may be used to improve the market segmentation strategy. The purpose of this study is to present an approach that indicates how the technique may be used to analyze services. The organization that is used from the standpoint of service analysis is the commercial bank.

The contribution of the paper is twofold: the focus on services in marketing management, and the unique delineation of the ideal point (I). Previous papers on the subject of perceptual mapping have primarily dealt with product attributes. This study, however, will essentially deal with product services. Researchers have interpreted the ideal point to represent a position in an n-dimensional space which best describes the mean preferences of the respondents (Neidell, 1969). As such it was desirable to be in a juxtaposition with this point. The closer the proximity, the more a product was described as fulfilling the needs of the consumer. In this study, the ideal point is based on service saliency. Therefore, the ideal point may be considered as the minimum acceptable amount of services. While in the normal context it is detrimental to product selection if a product possesses more of an attribute than signified by the ideal point, in this application, only lesser amounts of services hamper selection.

When making marketing decisions assumptions are formulated regarding the perceptions of bank customers. Management attempts to consider how customers perceive checking and loan policies, interest rates, hours of operation, branch office location, the cooperativeness of the personnel and so on. However, little is really known about just how present and potential customers perceive a given service or offering of a bank (White and Woodside, 1973). The risk factor inherent in such interpretations leads to uncertainty regarding the payoff of various marketing strategies. This study extends the research and indicates how perceptions may be analyzed to optimize a set of decisions.

## The Data

The first data collection sequence of the study involved approximately 100 marketing students at Syracuse University. They were requested to respond to questions concerning those attributes that are important in their choice of a bank, which banks they frequent, and which banks are familiar by name. From the lists they supplied, 6 local banks and 11 bank attributes were selected. An additional restriction was later imposed by the researchers on the choice of banks; they must be full service and advertise their products. This additional limitation reduced the number of test banks to 4. The list of attributes selected by the respondents was as follows:

- a - Free Checking Service
- b - Efficient Service
- c - Bank Statement Service
- d - Loan Policy
- e - Fast Service
- f - Long Bank Hours
- g - Interest Rate Level
- h - Friendly Personnel
- i - Drive-in Service
- j - Bank-by-Mail
- k - Convenient Location

In the second data collection sequence, sixty subjects (40 marketing students at Syracuse University, 20 non-students selected on a convenience basis) were provided with two questionnaires: one was concerned with the saliency of bank services and the other involved perceptions of bank services. Initially, there was concern that the students would have perceptions radically different from the non-students, but a comparison proved this unwarranted. Both groups were combined in the study.

The saliency of bank services was measured via the paired comparison technique. Subjects were presented with a pair of bank services such as: long bank hours and fast service. The respondents were then asked to indicate which one of the pair was most important when conducting business with a bank. Each respondent circled either the first service in the pair or the second. A total of fifty-five paired comparison questions were involved.

The questionnaire on bank service perceptions required the respondents to rate each bank in terms of the 11 selected attributes. Respondents rated each attribute on an eleven point scale from "does not describe the bank" to "does describe the bank" (Assael, 1971). A space was also provided to allow a no response with the stipulation that this would result only if the individual lacked any opinion of the bank concerning that particular attribute. The subjects were informed that this opinion did not require personal experience with the bank.

## Methodology

The output from the survey on the saliency of bank services was analyzed via factor analysis and the Varimax routine. The 11 bank services were reduced to four factors. A correlation value of  $\pm 40\%$  was assigned as the cut-off for each variable; i.e., any variable whose correlation with a factor was greater than  $\pm 40\%$  was described as comprising the basis of that factor. Four factor names were assigned. They are:

- Factor 1 Internal Convenience (var. c, i, j, k)
- Factor 2 Operational Compatibility ( var. a, b, g)
- Factor 3 Non-service Compatibility (var. d, f, k)
- Factor 4 Convenience Services (var. e, h, k)

The eigenvalue in the factor derivation was stopped at 1.00. At this level four factors emerged that accounted for 58.2% of the total variance. If the eigenvalue had been reduced to the .88 level, the number of factors would have been increased to six and 75.1% of the variance would have been explained. A trade-off is involved here between variance accountability and graphic interpretability. It was decided to select four factors with six graphs rather than six factors that would have involved 15 charts in order to make the data more manageable.

Each respondent was positioned in an n-dimensional map on the basis of his factor scores (Nie, Bent and Hull, 1970).

The results of the survey on the perception of bank services was plotted in the same perceptual space as the service saliency data. This was accomplished by using scores derived in the factor analysis as a weighting mechanism. A zero response was eliminated from the calculation.

## Results

The results of the study are depicted in Figures 1 through 6. The symbols A, B, C, and D represent the positioning of the mean values of the individuals' perceptions of the four banks regarding each pair of factors. In other words, A, B, C, and D refer to bank service perceptions. The letter I in the graphs pertains to the mean value of the respondents in terms of the saliency of bank services. The graphs indicate on the basis of mean values the position of the banks in terms of what they are perceived to provide in the way of services and the importance customers place on various services when dealing with a bank.

## Marketing Strategy Implications

From the standpoint of marketing strategy, a number of implications exist. Three are: service re-positioning, market segmentation, and service surveillance.

In terms of service re-positioning one may observe (Figure 1-6) the general disparity between what a bank is perceived to provide and what services are important to the customer. In all but one figure the value I is in a different quadrant than the perceptions A, B, C, and D. This means in managerial terms that the banks are perceived as offering a package of services at a certain level and that these services are not congruent with the level desired

Figure 1

Perceptual Map: Internal Convenience  
and Operational Compatibility

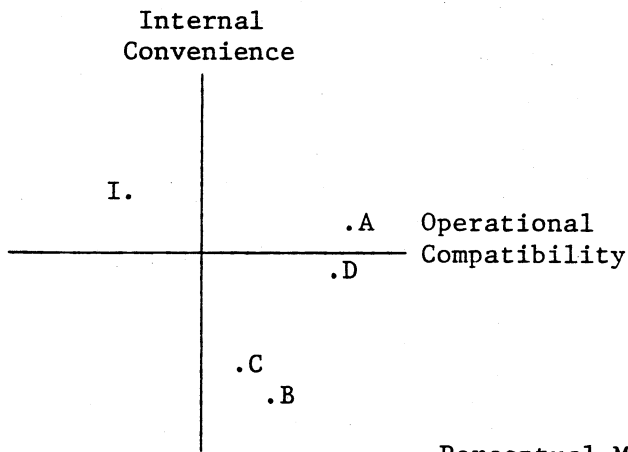


Figure 2

Perceptual Map: Internal Convenience  
and Service Compatibility

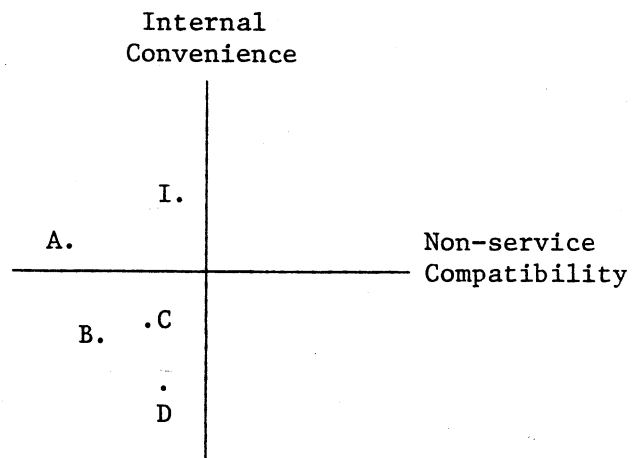


Figure 3

Perceptual Map: Internal Convenience  
and Convenience Services

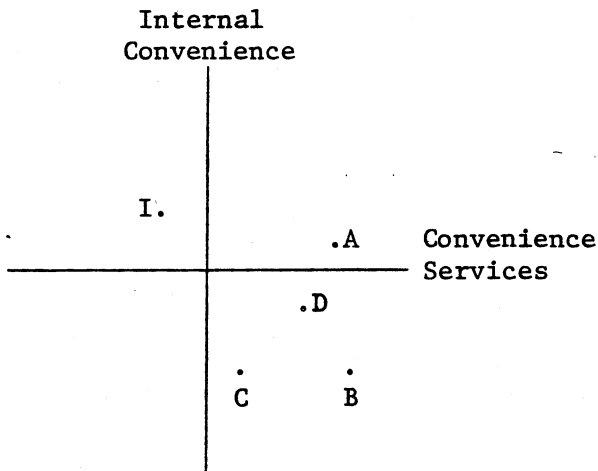


Figure 4

Perceptual Map: Operational Compatibility and Non-service Compatibility

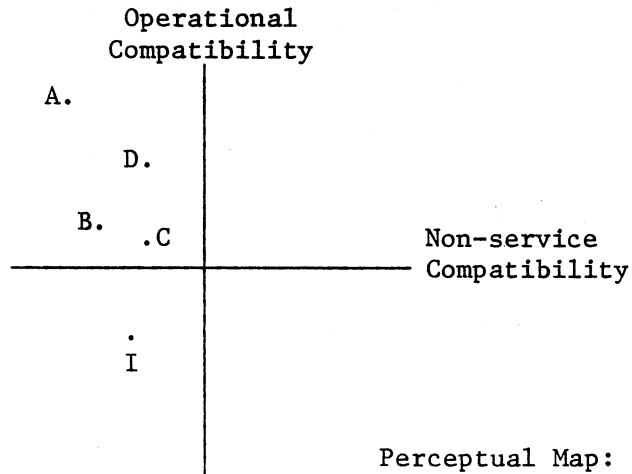


Figure 5

Perceptual Map: Operational Compatibility and Convenience Services

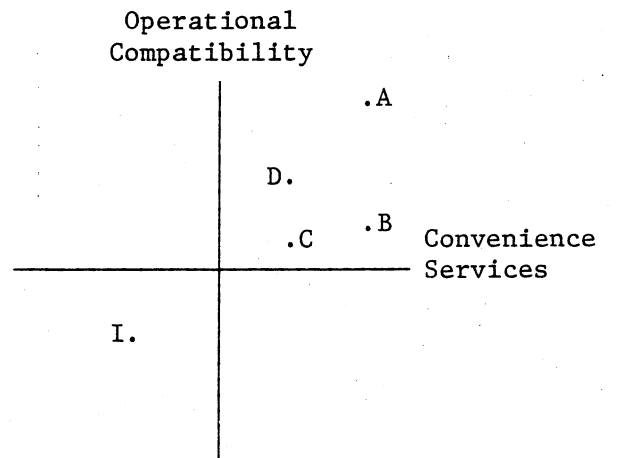


Figure 6

Perceptual Map: Non-service Compatibility and Convenience Services

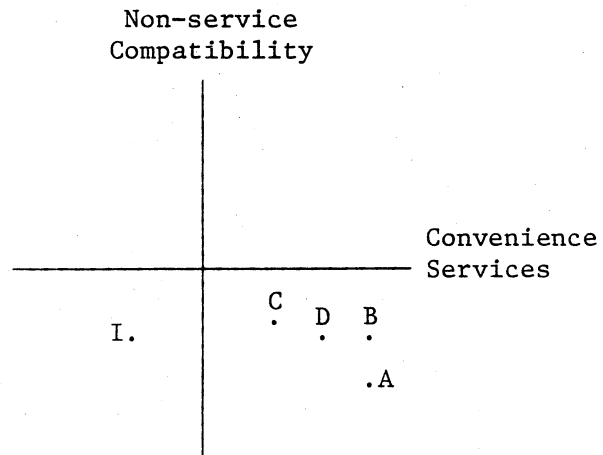


Figure 7

Perceptual Map of Mean Bank Locations  
and Actual Consumer Positions Relative to  
Internal Convenience and Non-service Compatibility

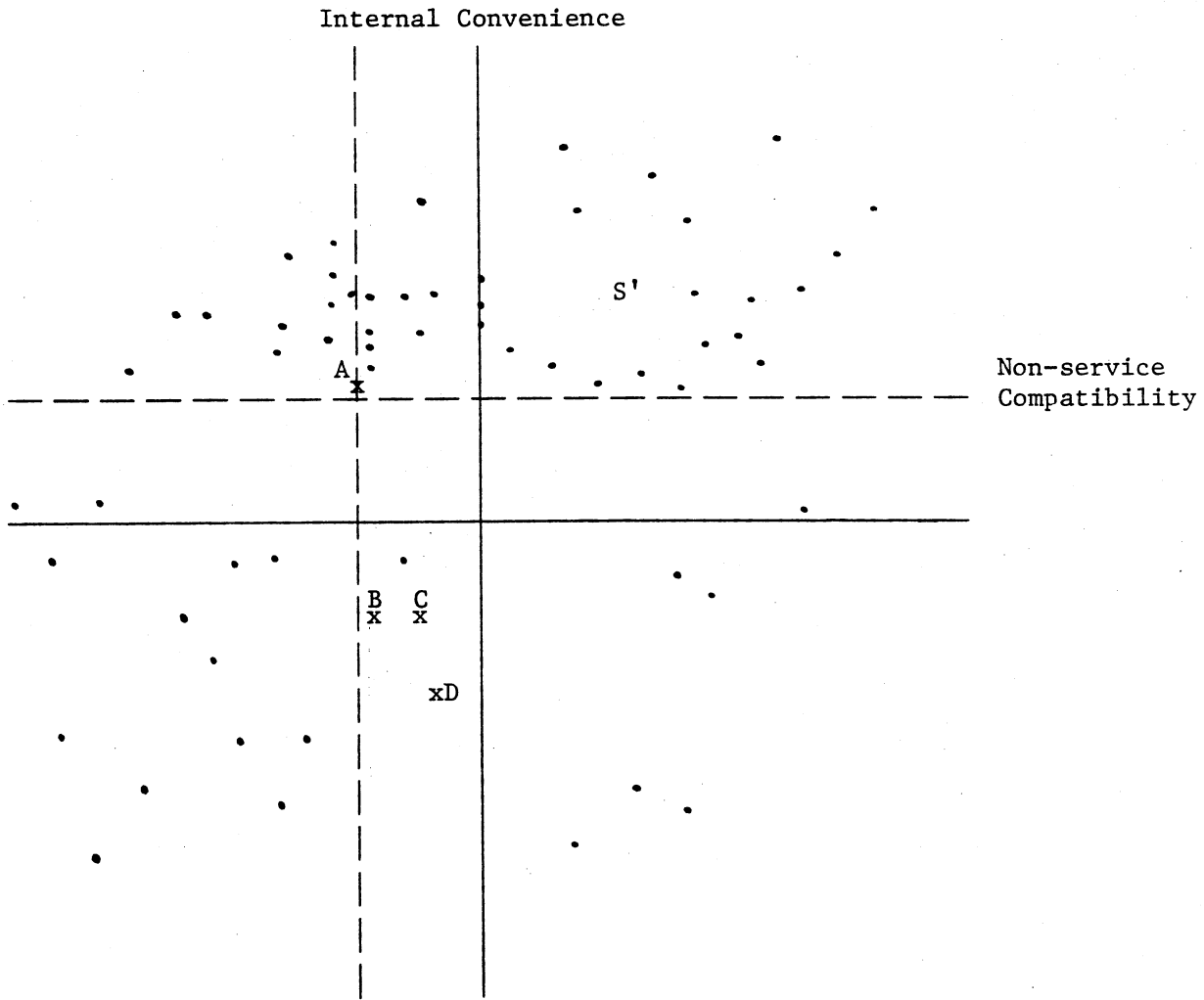
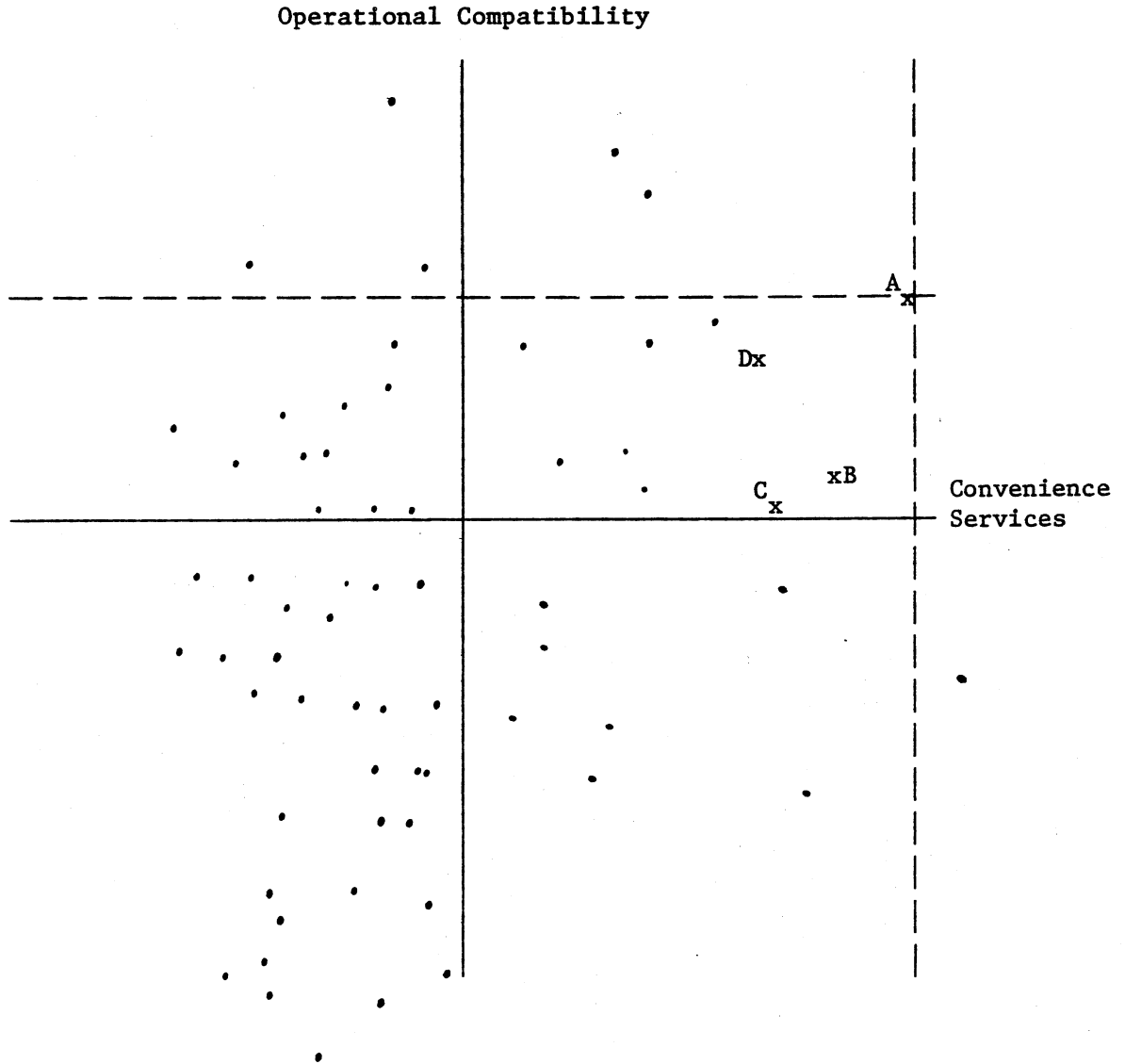


Figure 8

Perceptual Map of Mean Bank Locations  
and Actual Consumer Positions Relative to  
Operational Compatibility and Convenience Services



by the user. In other words, one may say that perceived service level does not match perceived service importance. The results of the study indicate that it would probably be advantageous for the banks to re-evaluate the nature of their services and through marketing communications bring them more in line with the evaluations of the customer. In addition, it may be noted that the banks form a cluster in varying degrees regarding their perceived services. This indicates that little service differentiation is viewed to exist among banks. The results suggest that a differentiation policy may be profitable.

The results of the study indicate that market segmentation may be used to



advantage. For example, Figure 7 and Figure 8 illustrate the location of bank service perceptions and bank service salience (points). The salience points in Figures 7 and 8 are useful as they identify the distribution of the perception of each user within the market sample regarding service salience. Using Figures 7 and 8 as examples, one may transpose the location of the x and y axes to a particular bank such as Bank A in this example. In analyzing Figure 7, it can be observed that 66% ( $40/60 \times 100$ ) of the respondents would like a bank to possess more internal convenience than Bank A is viewed to have by the sample. In addition, 61% ( $37/60 \times 100$ ) of the respondents would not be satisfied with the degree of non-service compatibility that the bank is perceived to have. Figure 8 indicates that in terms of convenience services and operational compatibility, 3% ( $1/60 \times 100$ ) and 8% ( $5/60 \times 100$ ), respectively, of the respondents will be dissatisfied.

From a managerial standpoint, the above statistics may be interpreted as follows: Bank A appears to be relatively weak in non-service compatibility and internal conveniences as indicated by the 61% and 66%. It would be in these two areas that Bank A has the most to gain from marketing expenditures. As shown in Figure 8, Bank A is depicted to be essentially right on target relative to its efforts in convenience services and operational compatibility. Additional effort in these areas would not appreciably increase overall market satisfaction.

A question that arises is: to what consumer cross-section does the bank want to appeal--the total market, or some particular segment. When considering the total market, the normal approach is to position oneself in a juxtaposition with the mean ideal point. However, with the set of factors utilized in this case, the best place to be, as shown in Figure 7, is in the upper right hand corner of the upper right hand quadrant (again, there must be recommended caution on the issue of assessing strength of factors as desirability of the service). As the position of the bank moves down and away from this point, fewer and fewer consumers feel that the bank can satisfy its needs. Because of financial constraints, it is impractical to attempt the satisfying of all consumers, therefore, a trade-off is required. A bank's best alternative is to move into a position of competitive advantage over the other banks in the area. Considering Figure 7, if Bank A were able to move its service mix to Point S', not only would it be satisfying a good percentage of the market's needs, but also would be in a closer proximity to the remainder of the market than the other banks.

Perceptual mapping also offers management information that may be used in service surveillance. An examination of perceptual maps compiled at different time intervals will show the changes in consumer opinion of the bank. It will also disclose any shift in consumer needs. This feature of perceptual mapping is quite beneficial in determining and gauging the effectiveness of promotional expenditures. However, it may be unwise for a bank to change the ratio of expenditures by factor (service) without knowing first the effect one factor has upon another. This can be achieved by changing expenditures on one factor while holding the other constant.

#### Summary

Perceptual mapping is a useful tool for analyzing the management of services. It is a technique that is not without its limitations, however. It can provide useful information to assist in service re-positioning, market segmentation, and service surveillance. From the standpoint of planning, perceptual mapping may indicate where effort should be concentrated or reduced.

In terms of bank marketing, this pilot study indicated that services are perceived to be relatively homogeneous among banks. In addition, it was found that the service level offered by banks was not in most cases synchronized with the service level mix believed to be important by the customer. Suggestions were made as to how this situation could be improved via three marketing management approaches.

#### FOOTNOTES

1. This study was supported in part by funds from the Manufacturers and Traders Trust Company. The authors also wish to thank Dr. N. Dobashi of the Psychological Services and Research Center at Syracuse University for his assistance and comments.
2. Graduate Student, School of Management, Syracuse University.
3. Associate Professor, Department of Marketing, School of Management, Syracuse University.

#### REFERENCES

- Assael, H. Perceptual mapping to re-position brands. Journal of Advertising Research, 1971, 2, 39-42.
- Green, P. E., & Carmone, F. J. Multidimensional scaling: An introduction and comparison of non-metric unfolding techniques. Journal of Marketing Research, 1969, 6, 330-341.
- Johnson, R. M. Market segmentation: A strategic management tool. Journal of Marketing Research, 1971, 8, 13-18.
- Lehmann, D. R. Judged similarity and brand-switching data as similarity measures. Journal of Marketing Research, 1972, 9, 331-334.
- Neidell, L. A. The use of non-metric multidimensional scaling in marketing analysis. Journal of Marketing, 1969, 33 (4), 37-43.
- Nie, N. H., Bent, D. H., & Hull, C. H. Statistical package for the social sciences. New York: McGraw-Hill, 1970.
- Wells, W. D., & Sheth, J. N. Factor analysis in marketing research. In R. Ferber (Ed.), Handbook of market research. New York: McGraw-Hill, 1974.
- White, S. D., & Woodside, A. G. Marketing in banking: Philosophies and actions. Journal of Bank Research, 1973, 265.



APPLICATION OF DISCRIMINANT ANALYSIS IN FORMULATING  
PROMOTIONAL STRATEGY FOR BANK CREDIT CARDS

James B. Wiley and Lawrence M. Richard  
Wayne State University

The objectives of this research are threefold: to present the complete analysis of psychographic data as suggested by Wells and Tigert, that is to go beyond reporting percentage figures on item responses; to extend the current research methodology by using generated factor scores as input to discriminant analysis for purposes of predicting user categories; and to examine the extent of concordance between this research and previously cited efforts. Factor analysis is used with linear discriminant analysis based on factor scores. The results suggest the following: Psychographics effectively can predict usage patterns for bank credit card users. As does multiple regression analysis, multiple discriminant analysis offers a diagnostic capability. It can serve as a guide to answering such questions as which set of variables are "important" from the standpoint of prediction. Finally, the results are suggestive of the explanatory potential of psychographic variables incorporated in predictive models.

Since 1963, marketing researchers have been exploring various market related measures of consumers' activities, interests, and opinions. These explorations represent what Edgar Pessemier and Albert Bruno have described as "the conscious attempts of researchers to expand the set of descriptors of consumer characteristics to fill the existent void between the economists' demographics and the psychologists' personality inventories" [Pessemier and Bruno, 1971].

Generally referred to as "life style research," such efforts have been conducted across a wide range of consumer activities. The most notable work has been carried out by Joseph Plummer [1971b, 1974], William Wells and Douglas Tigert [1971], Tigert [1966, 1969], Pessemier, F. S. DeBruicker, and Thomas Hustad [1968], Bruno [1971], and Pessemier and Bruno [1971].

Life style research most recently has been applied in the bank services area. Plummer [1971a] presented a psychographic profile of the users of bank credit cards, while Albert Pool [1974] extended Plummer's methodology somewhat through the application of discriminant analysis in examining cash dispensing machine users. This article represents an extension of both of these previous works.

#### Objectives of the Research

One objective of this paper is to examine the extent of concordance between this research and previously cited efforts [1971a]. The primary objectives,

however, are (1) to present a complete analysis of psychographic data as suggested by Wells and Tigert [1971], that is, to go beyond reporting percentage figures on item responses, and (2) to extend the current research methodology by using generated factor scores as input to discriminant analysis for purposes of prediction. Pool [1974] attempted to move in this direction, but he used only a limited number of psychographic items. This restriction was probably a function of instrument length prior to inclusion of the psychographic items. This study includes 53 items selected from several available item banks.

We believed it would be more meaningful for decision making to categorize users according to level of usage, rather than in a purely dichotomous fashion. For this reason, we have avoided the user versus nonuser categorization and have chosen to examine heavy versus light patterns. Accordingly, each card holder is classified as either (1) a "heavy" user (more than once a week), (2) a "medium" user (once a fortnight to once a week), or (3) a "light" user (less than once every two weeks). These categories correspond to usage categories adopted by the firm in related research.

#### Sources of Data

The respondents participating in this study were randomly selected from a list of applicants for cash dispensing machine cards. This sample is to be utilized in a study of cash dispensing machine as well as bank credit card users, but the results here pertain only to the latter category.

Standard demographics (Table 1) in conjunction with selected psychographic items (Table 2) were gathered. The latter were selected from previously utilized items [Plummer, 1971b; Pool, 1974] which have demonstrated reliability over time and stability across populations [Pessemier and Bruno, 1971]. One thousand applicants were sampled, with 35 percent responding to the questionnaire. Of the 350 respondents, 233 held bank credit cards.

Table 1 presents the demographic description of the credit card users surveyed. The heavy users, as opposed to light users, both male and female, tend to be younger (25-34), relatively highly educated (college graduate), employed in managerial and/or professional occupations, and have higher mean incomes. These results tend to conform to previous research [1971a]. Our objectives here are to go beyond such descriptive characterization of the data. A sequential analysis was adopted for this purpose.

#### Methods of Analysis and Results

The data are analyzed using factor and linear discriminant analysis. Results of the former are based on all respondents. Results for credit card usage, of course, refer to the sample of card holders. This latter sample proved to be sufficiently large for the "best" stepwise discriminant model to yield encouragingly stable results under double cross-validation procedures.

In utilizing factor analysis, the task is to reduce the large number of explanatory variables (AIO items) to a smaller number of presumably underlying variables or factors which then can be used as independent variables in subsequent analysis. This study follows a method proposed by Kendall [1957] which bases analysis upon the principal components of R, the correlation matrix of the total pool of AIO items. The components then summarize the "shared" variance in the explanatory pool and, by definition, provide an orthogonal set of predictors for subsequent analysis. The basic factor analysis model is

$$Z_i = a_{i1} + \dots + a_{ij} F_j + a_{iK} F_K + e_i \quad (1)$$

where

$Z_i$  = variable  $i$  in standardized form;

$F_j$  = hypothetical factor  $j$ ;

$a_{ij}$  = standardized multiple regression coefficient  
of variable  $i$  on factor  $j$  (factor loading);

$K$  = number of factors; and

$e_i$  = error term.

TABLE 1  
Demographic Summary

	<u>Male</u>				<u>Total</u>	<u>Female</u>				<u>Total</u>
	<u>Heavy Use</u>		<u>Light Use</u>			<u>Heavy Use</u>		<u>Light Use</u>		
<u>Age</u>	<u>#</u>	<u>Z</u>	<u>#</u>	<u>Z</u>		<u>#</u>	<u>Z</u>	<u>#</u>	<u>Z</u>	
Less than 25	7	13.2	8	7.7	15	4	23.5	12	21.1	16
25 - 34	28	52.8	37	35.6	65	8	47.1	18	31.6	26
35 - 44	3	5.7	23	22.1	26	1	5.9	3	22.8	4
45 - 54	9	17.0	26	25.0	35	2	11.8	8	14.0	10
55 over	6	11.3	10	9.6	16	2	11.8	6	10.5	8
<u>Education</u>										
Less than High School	1	1.9	1	1.0	2	0	0.0	1	1.8	1
High School	4	7.5	17	16.3	21	1	5.9	14	24.6	15
Some College	10	18.9	32	30.8	42	11	64.7	28	49.1	39
College Graduate	38	71.7	54	51.9	92	5	29.4	14	24.6	19
<u>Income</u>										
Less than 4999	0	0.0	1	1.0	1	2	11.8	2	3.5	4
5000 - 7999	3	5.7	4	3.8	7	2	11.8	7	12.3	9
8000 - 9999	2	3.8	1	1.0	3	3	17.6	9	15.1	12
10000 - 14999	13	24.5	34	32.7	47	4	23.5	28	49.1	32
15000 over	35	66.0	64	61.5	99	6	35.3	10	17.5	16
<u>Occupation</u>										
Craftsman	5	9.4	7	6.7	12	0	0.0	1	1.8	1
Clerical, Sales	2	3.8	12	11.5	14	6	35.3	16	28.1	22
Managerial	19	35.8	26	25.0	45	3	17.6	7	12.3	10
Professional	18	34.0	35	33.7	53	6	35.3	15	26.3	21
Others	9	17.0	24	23.1	33	2	11.8	17	29.8	19

\* N = 233, Total Male = 157, Total Female = 74, several categories total less due to missing data.

The principle factor loading matrix, A, for the ten principle components with eigenvalues greater than 1.00 was calculated and rotated according to the varimax criterion across all psychographic items. An element of this matrix,  $a_{ij}$ , indicates the correlation between attribute  $Z_i$  and factor  $F_j$ . The varimax rotation tends to maximize the correlation between an attribute and a single factor. An abbreviated version of A is presented in Table 2. The number and nature of factors recovered and the patterns of factor loadings appear to correspond closely with the experience of previous research. Many of the common variables identified here parallel those identified in Plummers' work [1971a] as well as in the more general study of Pool [1974]. These variables also closely correspond to those described by Pessemier and Bruno as exhibiting con-

TABLE 2  
Results of Factor Analysis

	Common Variables	Factor Loading	Common Variables	Factor Loading	Common Variables	Factor Loading
1.13	Dynamic Leader (10.1)* I think I have more self-confidence than most people.	.5881	Fashion Conscious (27.2) I usually have one or more outfits that are the very latest style.	.5368	Financially Satisfied (40.9) Our family income is high enough to satisfy nearly all our important desires.	.5173
1.14	I am more independent than most people.	.5740	If I must choose between the two I usually dress for fashion, not for comfort.	.6448	No matter how fast our income goes up we never seem to get ahead.	.4844
1.15	I think I have a lot of personal ability.	.6838	An important part of my life and activities is dressing smartly.	.6392	Unexpected situations often catch me without enough money in my pocket.	.6406
1.16	I like to be considered a leader.	.6698	Aesthetic Taste (31.1) I would like to spend a year in London or Paris.	.6113	I am always careful to carry plenty of money with me at all times.	.4141
1.17	My friends or neighbors often come to me for advice.	.5929	I would like to take a trip around the world.	.6128	Innovator (43.1) When I see a new brand on the shelf I often buy it just to see what it is like.	.7380
1.18	I sometimes influence what my friends buy.	.5058	I enjoy going through an art gallery.	.7022	I often try new brands before my friends and neighbors do.	.7237
1.19	People come to me more often than I go to them for information about brands.	.4369	I enjoy going to concerts.	.7505	I like to try new and different things.	.4559
2.9	Sports Minded (17.8) I like to watch or listen to baseball or football games.	.8145	I like ballet.	.5550	Information Exchanger (45.3) I often seek out the advice of my friends regarding which brand to buy.	.5093
2.10	I usually read the sports page in the daily paper.	.8190	Credit User (34.6) I buy many things with a credit card.	.6685	I spend a lot of time talking with my friends about products and brands.	.6210
2.11	I thoroughly enjoy conversations about sports.	.8738	I like to pay cash for everything I buy.	.7501	My neighbors or friends usually give me good advice on what brands to buy in the grocery store.	.6880
2.12	I would rather go to a sporting event than a dance.	.7960	It is good to have charge accounts. To buy anything, other than a house or a car, on credit is unwise.	.5836		
3.36	Price Conscious (22.8) I usually shop a lot for "specials."	.8343	Community Minded (37.8) I am an active member of more than one service organization.	.7923		
3.37	I usually watch the advertisements for announcements of sales.	.8376	I do volunteer work for a hospital or service organization on a fairly regular basis.	.7914		
3.38	A person can save a lot of money by shopping around for bargains.	.6894	I like to work on community projects.	.7500		

Note: Derived factor names from previous studies by Plummer.

Note: In every case the loading reported is the single factor largest for the item.

\*Cumulative explained variance.

siderable stability and reliability over several previous studies. The factor names in Table 2 are the same as those which appeared in these previous studies.

As a result of the factor analysis, respondents are characterized not only by a demographic profile, but also by a profile of scores that presumably summarize their attitudes, interests, and opinions in selected areas. Furthermore, each card holder can be classified as (1) a "heavy" user (more than once a week), (2) a "medium" user (once a fortnight to once a week), or (3) a "light" user (less than once every two weeks). Can this information be used to predict card holders' usage patterns? If so, which type of information is most useful?

Multiple discriminant analysis can be used to answer such questions.<sup>1</sup> The basic idea is simple; linear combinations of variables are sought that will maximize differences between the classes relative to differences within the classes. In order to do this, we calculate a vector of weights,  $w_1$ , for each of the classifications entering analysis:

$$w_1 = V^{-1}m_1 \quad (2)$$

where

$w_1$  = a vector of weights;

$V$  = pooled dispersion matrix; and

$m_1$  = mean scores for each class,  $1 = 1, K$ .<sup>2</sup>

In addition, we calculate a constant for each class,  $c_1$ :

$$c_1 = m_1' \cdot w_1 / 2. \quad (3)$$

Let us suppose we wish to decide which usage pattern characterizes a given card holder and that the card holder is characterized by the vector of scores  $S_i$ . In order to do this, we compute a score for class:

$$S_{i1} = S_i' \cdot w_1 - c_1. \quad (4)$$

We assign the individual to the class for which he obtains the highest score.

This procedure assumes the a priori probability of a card holder falling into a class (usage pattern) is equal to  $1/K$  for every class. If we can assign differential probabilities to classes, we can refine the classification procedure. We add to a person's score for each class the value  $\log_e P_1$ , where  $P_1$  is the a priori likelihood of a card holder having adopted the 1(th) usage pattern. That is

$$S(\text{adj})_{i1} = S_{i1} + \log_e P_1. \quad (5)$$

The result of applying a variety of multiple linear discriminant functions to card holders characterized in terms of psychographic and demographic variables is summarized in Tables 3, 4, and 5. Results for four models are reported. The first consists of the "best" stepwise model. That is, predictions



TABLE 3

Two-Way Classification  
Stated (Actual) and Pre-  
dicted (Classified) Credit  
Card Usage; Four Models

	Classified			Actual		
	(1)	(2)	(3)	(1)	(2)	(3)
Total	35	101	97			
Best <sup>1</sup>						
Factor <sup>2</sup>	25	28	13			
Demographic <sup>3</sup>	24	25	14			
All <sup>4</sup>	21	27	30			
Expected <sup>5</sup>	28	10	9			
	(11.5)	(33.5)	(32.0)			
Best	6	47	27			
Factor	6	54	27			
Demographic	10	57	44			
All	4	72	16			
Expected	(11.5)	(33.5)	(32.0)			
Best	4	25	57			
Factor	5	22	56			
Demographic	4	16	23			
All	3	18	72			
Expected	(11.5)	(33.5)	(32.0)			

<sup>1</sup> Best stepwise model; factors 1, 3, 4, 6, 7, 8 (Table 2) and income (Table 1), p to enter .05. The same variables were selected under forward and backward selection procedures. Chi-square (classified/actual) = 61.74, df = 4. Contingency coefficient = .45.  
<sup>2</sup> Factor scores only model, factors listed in Table 2. Chi-square = 52.95, df = 4. Contingency coefficient = .43.  
<sup>3</sup> Demographics only model, demographics listed in Table 1. Chi-square = 16.05, df = 4. Contingency coefficient = .25.  
<sup>4</sup> All AIO items (Table 2) plus demographics (Table 1). Chi-square = 168.20, df = 4. Contingency coefficient = .64.  
<sup>5</sup> Expected number in each classification with equal a priori assignment to categories. Expected chi-square = 0.0.

TABLE 4

Double Cross Validation  
Equal Prior Probabilities

	Classified			Actual		
	(1)	(2)	(3)	(1)	(2)	(3)
Total	35	101	97			
Best <sup>1</sup>						
Factor <sup>2</sup>	21	22	13			
Demographic <sup>3</sup>	17	27	17			
All <sup>4</sup>	11	30	32			
Expected <sup>5</sup>	15	21	18			
	(11.5)	(33.5)	(32.0)			
Best	10	52	29			
Factor	12	47	29			
Demographic	14	49	42			
All	14	37	27			
Expected	(11.5)	(33.5)	(32.0)			
Best	4	26	55			
Factor	6	27	54			
Demographic	10	21	23			
All	6	33	52			
Expected	(11.5)	(33.5)	(32.0)			

<sup>1</sup> Best stepwise model; factors 1, 3, 4, 6, 7, 8 (Table 2) and income (Table 1), p to enter .05. Chi-square (classified/actual) = 50.61, df = 42.  
<sup>2</sup> Factor scores only model, factors listed in Table 2. Chi-square = 30.42, df = 4. Contingency coefficient = .34.  
<sup>3</sup> Demographics only model, demographics listed in Table 1. Chi-square = 1.36, df = 4. Contingency coefficient = .071.  
<sup>4</sup> All AIO items (Table 2) plus demographics (Table 1). Chi-square = 18.30, df = 4. Contingency coefficient = .27.  
<sup>5</sup> Expected number in each classification with equal a priori assignment to categories. Expected chi-square = 0.0.

TABLE 5  
Double Cross Validation  
Proportional Prior Probabilities

	Classified	Actual		
		(1)	(2)	(3)
Total		35	101	77
Best <sup>1</sup>	(1)	8	8	0
Factor <sup>2</sup>		9	8	2
Demographic <sup>3</sup>		0	2	3
All <sup>4</sup>		14	28	16
Expected <sup>5</sup>	(5)	(5)	(15)	(14)
Best	(2)	22	68	41
Factor		20	67	39
Demographic		19	61	51
All		15	38	28
Expected	(15)	(15)	(44)	(41)
Best	(3)	5	24	54
Factor		6	26	55
Demographic		16	37	43
All		6	34	53
Expected	(14.5)	(14.5)	(42)	(38)

<sup>1</sup>Best stepwise model, factors 1, 3, 4, 6, 7, 8 (Table 2) and income (Table 1), p to enter .05. Chi-square (classified/actual) = 39.54, df = 4. Contingency coefficient = .380.

<sup>2</sup>Factor scores only model, factors listed in Table 2. Chi-square = 38.94, df = 4. Contingency coefficient = .378.

<sup>3</sup>Demographics only model, demographics listed in Table 1. Chi-square = 2.64, df = 4. Contingency coefficient = .106.

<sup>4</sup>All AIO items (Table 2) plus demographics (Table 1). Chi-square = 18.88, df = 4. Contingency coefficient = .274.

<sup>5</sup>Expected number in each classification with a priori assignment to categories proportional to sample frequency. Expected chi-square = 0.0.

are based on variables selected through stepwise procedures. Psychographic factor scores provide the input for the second model, demographic variables for the third model.<sup>3</sup> The input for the fourth model consisted of all (53) AIO items plus all demographics. Table 3 results using discriminant functions derived from scores of the classified individuals. Tables 4 and 5 report results using discriminant functions derived from scores of one half of the sample to classify the other half, and vice versa. Comparison of the results in Table 3 with those in Tables 4 or 5 shows the considerable "shrinkage" that can occur when discrimination functions derived from one sample are used to classify individuals drawn from another sample.

#### Results, Summary, and Conclusions

In reviewing the results of the analyses, several points should be made. Our first objective was to report the complete results of a factor analysis of psychographic items. As shown in Table 2, our findings tend to be supportive of previous ones. The common variables in this analysis have tended to appear in previous works [Plummer, 1971a; Pool, 1974], implying that they do exhibit relative stability and reliability over time [Pessemier and Bruno, 1971].

Our second objective was to extend the current methodology through the use

of discriminant analysis with generated factor scores being used as input. Table 3 presents the two-way classification of observations according to usage patterns. More important, Tables 4 and 5 present double cross validations based on equal and proportional prior probabilities for assignment. An examination of these tables supports the following positions:

- (1) Psychographics effectively can predict usage patterns for bank credit card users. For example, when using the best stepwise model, 55 percent of the observation are correctly classified, which is better than the number expected by chance with equal a priori assignment (33 percent), better than with proportional a priori assignment (37 percent), and better than the number expected with assignment of everyone to the most frequent class (43 percent). The second criteria corresponds to Morrison's [1969]  $C_{pro}$  criteria. The latter corresponds to his  $C_{max}$  criteria.
- (2) Psychographics do a better job than demographics alone.
- (3) The critical variables that seem to do the best job of predicting usage categories are factors 1 (Dynamic Leader), 3 (Price Conscious), 4 (Fashion Conscious), 6 (Credit User), 7 (Community Minded), 8 (Financially Satisfied), and income.

Table 6 presents the mean scores for the factor scores on these factors, plus income. The table suggests that the observed income difference between heavy and light users of bank credit cards is a significant one.

It should be pointed out that due to the sample utilized, these results are not intended to be generalizable to the population of all bank credit card users. However, these data can and should be treated as inputs to further study as they illustrate how relatively stable consumer characteristics may be linked to an indicant of service usage. As Frank et. al [1972] point out, such links can serve a dual purpose:

1. To facilitate reaching the attitudinal segments via media
2. To better understand the nature of this segment and hence better design promotional messages aimed at its members

In more specific terms, these are data that can provide guidance for the promotional strategist who is concerned with constructing a tentative answer to the following questions: "Who are the users of these services, and how can the bank most effectively promote continued and/or expanded use of the services?" For example, these results (Table 6) suggest Heavy users tend to have significantly higher incomes. They also tend to exhibit significant differences in individual self-confidence; that is, they tend to be more competent than Light users. Heavy users of bank credit cards also tend to be both significantly more price and fashion conscious. They tend to be more community minded or community oriented. These individuals also tend to be financially satisfied. This is, at least intuitively, in line with the point that they fall in the high income categories.

TABLE 6

Summary One-Way Anova  
Selected Variables

<u>Variable</u>	<u>Mean</u>			<u>F</u>	<u>Significance</u>
	"LIGHT"	"MEDIUM"	"HEAVY"		
Income	4.18	4.13	4.60	3.07	.0480
Factor 1	5.91	6.01	6.48	5.33	.0050
Factor 3	2.71	2.24	2.14	6.50	.0020
Factor 4	2.82	3.27	3.46	7.24	.0009
Factor 6	4.83	5.16	5.80	14.11	.0000
Factor 7	-.32	-.00	-.21	2.12	.12
Factor 8	3.77	3.91	4.30	3.57	.0290

FOOTNOTES

1. An alternative approach could be based on canonical discriminant analysis. Use of factor scores, however, appears to offer the advantages of canonical analysis while retaining the explanatory potentiality of AIO analysis.
2. The analogy with the multiple regression equation

$$b = R^{-1} K$$

is evident. Both  $w$  (Equation 2) and  $b$  are vectors of weights. The vector of differences between group means of the variables ( $d$ ) is analogous to the vector of correlations between the predictor variables and the criterion variable ( $K$ ).

3. Sex and occupation were coded as dummy variables.

REFERENCES

- Bruno, Albert V. An empirical model for the evaluation of television vehicles. Unpublished doctoral dissertation, Krannert Graduate School of Industrial Administration, Purdue University, 1971
- Frank, R., Massy, W., and Wind, Yoram. Market segmentation. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1972.
- Kendall, M. G. A course in multivariate analysis. London: Charles Griffin and Company Limited, 1957.
- Morrison, Donald. On the interpretation of discriminant analysis. Journal of Marketing Research, 1969, 6, 156-163.
- Pessemier, Edgar, and Bruno, Albert. A empirical investigation of the reliability and stability of selected activity and attitude measures. In David Gardner (Ed.), Proceedings. College Park, Maryland: Association for Consumer Research, 1971, 389-402.

- Pessemier, Edgar, Hustad, Thomas P., and DeBruicker, F. S. The Purdue consumer behavior research project. Unpublished working paper, Krannert Graduate School of Industrial Administration, Purdue University, 1968.
- Plummer, Joseph T. Life style patterns and commercial bank credit card usage. Journal of Marketing, 1971, 35 (2), 35-42. (a)
- Plummer, Joseph T. Life style and advertising: Case studies. Paper presented at the 54th Annual International Marketing Congress, American Marketing Association, April, 1971. (b)
- Plummer, Joseph T. The concept and application of life style. Journal of Marketing, 1974, 38 (1), 33-37.
- Pool, Albert A. Application of discriminant analysis in formulation of promotional strategy for cash dispensing machines. Journal of Bank Research, Spring, 1974, 13-19.
- Tigert, Douglas J. Consumer typologies and market behavior. Unpublished doctoral dissertation, Krannert Graduate School of Industrial Administration, Purdue University, 1966.
- Tigert, Douglas J. Psychographics: A test-retest reliability analysis. Marketing Involvement in Society and the Economy. Proceedings of the American Marketing Association, August, 1969, 311-315.
- Wells, William D., and Tigert, Douglas J. Activities, interests and opinions. Journal of Advertising Research, 1971, 11 (2), 27-34.

DEVELOPMENT OF A SCALE FOR INNOVATIVENESS<sup>1</sup>

Clark Leavitt<sup>2</sup> and John Walton<sup>3</sup>  
The Ohio State University

This paper describes the construction of a scale to measure innovativeness--a psychological trait underlying adoption of new ideas, services and products. The study utilizes as a model a psychometric methodology that has demonstrable power in personality scale construction in order to demonstrate its potential for use in marketing research. Four steps were gone through in applying the model to innovativeness: (1) defining the trait, (2) creating a large pool of items based on that definition, (3) utilizing techniques for suppression of response bias to select the best items in terms of a validity/reliability criterion, (4) evaluation in terms of convergent and discriminant validity. A fifth criterion, predictive validity, awaits future developments.

This is a report on the first of a series of studies designed to understand the individual psychology underlying the adoption of new ideas and products. One of the determinants of time-of-adoption may be the individual's standing on some broad personality trait. We are concerned here with the construction of an instrument to measure such a trait.

The use of personality variables in marketing has not had a happy history. From the time of Evans' attempts to discriminate Ford and Chevy owners to the present, the record has been marked by failure, or, at best, limited success. Recently there has been renewed reason for encouragement as a result of attempts to relate patterns of personality traits with patterns of product choice (see Worthing et al., 1973, for reference). The interaction between traits of the person and other variables such as situational measures has also shown promise (Endler and Hunt, 1968).

The basically poor record should come as no surprise. As has been frequently pointed out (Mischel, 1968) by personality psychologists, global traits have very little predictive power in most areas. Beyond this, the choice of measures in marketing seems unnecessarily *ad hoc*, often involving instruments with poor or unknown technical specifications and with little thought about the nature of the product choice to be predicted.

Yet it seems premature to dismiss the possibility that a single global trait can be worth the candle in marketing, particularly in predicting time-of-adoption behavior, dependent as it is on information utilization of the kind often influenced by personality traits. Therefore, this study was done to find or create an appropriate trait measure.

#### The Adoption Process

Considerable research during the past decade has begun to highlight and quantify major elements of the diffusion process of new product and other

innovations. Everett Rogers (1962), in his now classic work, postulated a diffusion paradigm that was capable of generalization to diverse behavioral disciplines. Robertson (1971), using the Rogers paradigm as a point of departure, established the practical and theoretical efficacy of marketing innovation research.

The diffusion of various types of marketing innovations has been studied with the new line of ladies hats (King, 1964); instant coffee (Frank, Massy, Morrison, 1964); auto diagnostic centers (Engel, Kegerreis, Blackwell, 1970); a detergent (Pessemier, Burger, Tigert, 1967); and touchtone telephones (Robertson, 1967), among others.

A degree of commonality can be found in the research methodologies of most of these studies: (1) the dependent variable, adoption, is operationalized by the researcher; (2) an expedient assortment of independent variables is specified; (3) correlation or regression analysis is used to measure the degree or strength of the relationship between a variety of independent variables and the dependent variable.

The results from this paradigm have not been encouraging. Indeed, the typical finding has been that only a small amount of variance in the dependent variable is explained by a variety of demographic and personality variables used as predictors. Two alternative explanations for these data are possible. First, repeated negative findings may indicate that there is no relationship between adoption and independent consumer variables, especially at the micro or individual level. If this is the case, then research resources should be put to more fruitful uses. The second explanation is that the specification of the research variables is inadequate. Since the operationalizing of most demographic variables is fairly standard this criticism mainly concerns the personality variable.

It is difficult to accept the first explanation. Behavior adoption of innovations has been successfully studied in a variety of other disciplines (Rogers and Shoemaker, 1971). The application of this concept to marketing and consumer behavior has too much face validity to dismiss it easily.

The personality variables are the soft underbelly of the problem. It is generally recognized (see Kassarjian, 1971, or Worthing, et al., 1973) that selection of personality variables as such has not been carried out according to a systematic rationale for the appropriateness of the specific measure. In particular, the authors know of no studies that have devised specific measures for the behavior under consideration with anything approaching the care that could be used in arriving at paper-and-pencil measures of personality traits.

To test the effectiveness of the strategy of customized construction of a personality scale designed to fit a specific problem, the present project has attempted to measure innovativeness--a presumed trait underlying adoption of innovations.

This study was undertaken with the thought in mind that the explication of methodology of psychometric test construction is an area of opportunity for marketing. However, it is probably worth the considerable cost and effort only where the psychological trait has a potential significance extending beyond one product area. This was one reason for using adoption behavior. Along with the specificity of each of the content areas where adoption takes place, it is possible that the newness of the innovation as such will have a general effect over all categories. Also, instruments constructed by the technique

used here have been among the few to show some success in marketing (Worthing et al., 1973).

### Psychometric Methodology

There are two important strategy considerations in scale construction. First, the manner in which the items are generated: here the choice is between a rational approach versus an empirical approach. In the former the items represent the investigator's interpretation of an idea derived from theory or clinical intuition of a psychological trait that is real or important in a conceptual sense. The empirical strategy calls for the unsystematic selection of a large number of items taken from various sources which are then sorted on the basis of their ability to predict a criterion without regard to whether or not the relationship makes any sense beyond simple expediency.

Given a pool of items derived in either way, the second strategy concerns the manner in which the final items are selected. This may involve the issue of response bias as well as other means for maximizing reliability and validity.

A valid test is one that measures what we want to measure, all of what we want to measure, and nothing but what we want to measure (Thorndike and Hagen, 1961). The three principal types of validity include content validity, predictive validity, and construct validity.

Content validity refers to the "adequacy with which a specified domain of interest is sampled" (Nunnally, 1967). The establishment of content validity is intimately related to the choice of strategies mentioned above... a universe of content is delimited, and then items are systematically drawn from that universe. Frequently, judges or panels of experts are required to assure the adequacy of content validation procedures, especially with the rational approach.

While content validity is dependent to a large degree upon subjective considerations, predictive validity can be more objectively established. Predictive validity refers to the magnitude of the association between the scale score and the behavior or criterion that the scale is attempting to identify. To the extent that a high degree of association exists between the scale score and the criterion, then the scale is said to have high predictive validity.

Construct validity refers to the a priori meaningfulness of the variable or construct under investigation. More specifically, construct validation provides at least a partial answer to the question, Does this scale tell us something about the subjects under investigation that is broadly meaningful and significant? (Thorndike and Hagen, 1961).

When groups known to differ on the particular variable are easily identifiable, then assessment of construct validity is trivial.<sup>4</sup> However, when abstract constructs such as innovativeness are studied, then construct validation is more difficult. Campbell and Fiske (1959), as part of a more generic multi-trait/multi-method technique, suggest that convergent and discriminant validity may be one way to approach the assessment of construct validity. Using this approach, other measures of the same trait and sometimes other traits are hypothesized to be related positively (convergent) or be independent to an extent beyond that caused by common method variance. If these hypotheses are upheld empirically, then construct validity is affirmed.



The second important issue in scale construction is reliability. In this context, reliability has two distinct meanings. First, reliability measures the degree to which the items that make up the test are interrelated. This represents an internal consistency approach to reliability. Reliability has also been defined as the degree to which scores from a particular administration of a given test are related to scores on a subsequent administration of that same test to those same subjects. This represents a temporal stability approach. Both indices of reliability are important and should be demonstrable.

The final important issue in scale construction is the suppression of response biases. Response biases include any systematic variance in the scale score that is not attributable to the trait being measured. Two principal sources of this variance are the traits of social desirability and acquiescence. Social desirability measures the need for social approval reflected in the tendency of a subject to respond to items in a way consistent with approved social norms rather than their true feelings. Acquiescence is the tendency of subjects to agree with statements. Both of these sources of variance must be controlled.

#### A Rational Approach to Scale Construction

Jackson (1971) suggests the application of four principles of scale construction that consider each of these issues. These principles are enumerated below.

1. An explicit, theoretically-based definition of a particular trait is essential prior to attempts at measurement;
2. Careful empirical selection of items for homogeneity contributes substantially to refined measurement;
3. Suppression of response biases such as social desirability and acquiescence is best undertaken at the level of item selection and scale development; and
4. Both convergent and discriminant components of validity must be considered at every stage of scale development if the final scales are to possess these properties (Jackson, 1967; from PRF manual, p. 15).

First, the diffusion adoption literature was evaluated as background data to assist in the development of the range of variables that were necessary for subsequent evaluation of results.

The literature review showed that open-mindedness (Jacoby, 1971), leadership (Ostlund, 1972), self-confidence (Ostlund, 1972), venturesomeness (Robertson and Kennedy, 1968), empathy (Rogers and Stanfield, 1968), gregariousness (King and Sproles, 1973), and student activism (King and Sproles, 1973) were sometimes related to adoption.

Second, a long tradition of psychological research relating suggestibility and persuasibility to lowered self-esteem was considered. In a contrasting vein, a series of studies begun by Asch emphasized the independence and lack of conformity of subjects. These two research trends seemed to constitute two cells of a matrix that might be labeled "good independence" and "bad conformity." The other two cells seemed relatively empty: "bad independence" being

related, perhaps, to negativism or--more recently--to reactance. Only the very general area of problem solving seems related to "good conformity." This would represent the ability of the individual to pick up cues or to change set in order to use the resources of the situation creatively to solve a problem.

It seemed possible that a trait existed that underlay the intelligent, creative, selective use of communication for solving problems. In other words, being persuasible where it was adaptive for the recipient of the persuasive message. Surely we sometimes learn something of value from mass media and from opinion leaders. This led to defining a trait of "innovativeness" as follows:

A person high on the trait of innovativeness is open to new experiences and often goes out of his way to experience different and novel stimuli particularly of a meaningful sort (not just thrill-seeking). Most important, he tends to make constructive use of information received whether sought or accidentally encountered. He has a low threshold for recognizing the potential application of ideas he gets from others but does not apply suggestions mechanically. Rather, he has the ability to transform information for his own use. His involvement in his own enterprises and acts is such that he looks for ways to change and improve them. Above all, he is responsive to communication in a selective, constructive way when the message has a valid relevance to his activities. He is objective in his evaluation although occasionally naive.

Psychologists have attempted to measure rigidity, dogmatism, closed mindedness, submissiveness, and suggestibility, but there have been few, if any, attempts to scale the open-minded, constructive trait that we have in mind.

A group of three experts was assembled and after discussing the definition and type of person who was high on the trait, during several sessions over a period of two weeks produced 144 prospective positive innovativeness items. In addition, a twenty-item social desirability scale was included to measure desirability response bias (Crown and Marlowe, 1964). This total set of items was then administered to 300 female respondents.

The Jackson (1967) method of item analysis was used next to maximize reliable content variability in relation to social desirability response variance. For each item, the correlation between the scale item and the total desirability score was subtracted from the correlation between that scale item and the total scale score. The square root of the remainder is the proportion of true content (innovativeness) variance for any given item. Jackson calls this value the Differential Reliability Index (DRI). Obviously, the higher the DRI, the more the item is reflective of the substantive content (innovativeness) of the item. From this administration, items with DRI's of .39 or greater were retained. This amounted to 29 positive items from the original pool of 144.

While this procedure eliminated desirability response bias, acquiescence was still a problem. Therefore, a second questionnaire was constructed containing 196 items to deal with additional problems of scale development. These items included the 29 positive innovativeness items, 33 negative innovativeness items, 14 psychological subscales (94 items), and a 40-item new-product checklist (to be discussed in another report).

The 29 positive innovativeness items were retained from the previous analysis. The negative innovativeness items were included to alleviate acquiescence response bias. A large number of negative items were constructed by reversing the positive items in a sensible way. Some of these items were literal reversals of the positive items, while the remainder were more broadly a negation of innovativeness.

Fourteen psychological scales were included for construct validation purposes. If innovativeness is an independent trait it should show only moderate relationships with these other traits when compared to the correlation between the two innovativeness forms. Descriptions of these scales and hypothesized direction of relation with innovativeness are shown in Table 1. These hypothesized relationships were drawn from the sources previously cited. The questionnaire was administered to 299 women from a variety of demographic categories.

TABLE 1  
Description of Psychological Scales and  
Hypothesized Relationships with Innovativeness

Scale	Description	Hypothesized Relationship
Well being	Sense of well being about life	+
Political control	Internal control of political events	+
Internal/External	Internal control of personal events	+
Good spirits	Feeling in a happy mood	+
Confidence in others	Faith in abilities of others	+
Leadership	Functioning as head of a group	+
Culture vulture	Enjoyment of cultural activities	+
Inflated self-esteem	Extreme view of self-worth	-
Competence	Ability to perform specific tasks	+
Helplessness	Unable to control events	-
Energetic	Prone toward vigorous activity	+
Communication	Active processor of environmental stimulation	+
Community-minded	Concern for local affairs	+
Counter culture	Inclination toward non-traditional activities	+

Symbols: + Positive Relationship      - Negative Relationship

## Results

### Item Selection

Items selected for the final scales were given consideration under various criteria, primarily the whole-part correlation of each innovativeness positive item with positive total and each negative item with the negative item total. To qualify for inclusion in the scale, positive innovativeness items had to correlate at least moderately high with the positive total. In a similar manner negative items needed to be correlated at least moderately with the negative item total. The 20 positive and 20 negative items which most closely met these criteria were selected for the innovativeness scale. A listing of these items is available on request.<sup>5</sup>

Only scores of females are included in these data. Therefore, it is possible that these items are relevant only for women. A preliminary analysis

has been undertaken for male subjects and the similarity of the male data suggests that the items will work for both males and females.

### Reliability

Two internal consistency measures of reliability were calculated. These measures included the Spearman-Brown split-halves method and Kuder-Richardson formula 20. The value of these reliability coefficients for the 40-item test and alternate forms A and B are shown in Table 2.

TABLE 2  
Reliability Data

Method	40-Item Test	Form A	Form B
Spearman-Brown	.90	.84	.77
Kuder-Richardson, 20	.88	.80	.77

The reliability coefficients for the 40-item test are quite high. Corresponding reliability values for both alternate forms are, as expected, somewhat lower. These values represent, however, acceptable levels of reliability and justify the decision to develop two forms.

Divergent validity was assessed through the use of the fourteen psychological subscales. As shown in Table 1, it was hypothesized that innovativeness should show a moderate positive relation to political control, internal/external control, confidence in others, leadership, culture vulture, competence, energetic, communication, community-minded, counterculture, well-being, and good spirits.... a negative relationship was hypothesized with helplessness and inflated self-esteem. Actual data on these relationships are presented in Table 3.

The data in Table 3 indicate that these hypotheses were upheld with a few exceptions: internal/external control and communication were slightly negatively related to innovativeness, and inflated self-esteem was positively related to innovativeness. None of the correlations are high enough to suggest identity. These results indicate that at least a moderate degree of independence is evident for Innovativeness.

### Conclusion

The objective of this paper was to construct a valid and reliable scale to measure innovativeness--a trait postulated to underlie adoption behavior. Data presented indicate that either of two forms meet this objective. The success of these scales and, by inference, the developmental methodology outlined in this paper depend on the utility of these scales and this methodology for marketing researchers and practitioners in many different environments. The authors eagerly await these results. Meanwhile, further studies are underway to assess the effectiveness of Innovativeness in predicting adoption behavior of various kinds.

TABLE 3  
Divergent Validity Correlation of Psychological Scale  
Total Scores with Innovativeness Scales

Psychological Subscales	40-Item		
	Test	Form A	Form B
Well-Being	.22	.26	.16
Political Control	.25	.24	.25
Internal/External	-.08	-.13	-.02
Good Spirits	.30	.31	.28
Confidence in Others	.20	.20	.18
Leadership	.44	.36	.48
Culture Vulture	.46	.45	.43
Inflated Self-Esteem	.31	.28	.31
Competence	.32	.33	.33
Helplessness	-.35	-.39	-.32
Energetic	.04	.05	.05
Communication	-.10	-.15	-.08
Community-Minded	.05	.01	.09
Counter Culture	.39	.40	.40
Form A		---	.84
Form B		.84	---

#### FOOTNOTES

1. This research was supported by the Leo Burnett Company, Chicago.
2. Clark Leavitt is Professor of Marketing, The Ohio State University.
3. John Walton is a Ph.D. student at The Ohio State University.
4. See Crano and Brewer, 1973, p. 253, for a complete explication.
5. A copy of the final forms can be obtained by writing Prof. Clark Leavitt, The Ohio State University, 1775 College Road, Columbus, Ohio 43210

#### REFERENCES

- Asch, S. E. Social psychology. New York: Prentice-Hall, 1952.
- Bentler, P. M., Jackson, D. N. & Messick, S. Identification of content and style: A two-dimensional interpretation of acquiescence. Psychological Bulletin, 1971, 76, 186-204.

- Campbell, D. T. & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.
- Crano, W. D. & Brewer, M. B. Principles of research in social psychology. New York: McGraw-Hill, 1973.
- Crowne, Douglas P. & Marlowe, David. The approval motive: studies in evaluative dependence. New York: John Wiley and Sons, Inc., 1964.
- Endler, N. & Hunt, G. M. B. S-R inventories of hostility and comparisons of the proportion of variance from persons, responses, and situations for hostility and anxiousness. Journal of Personality and Social Psychology, 1968, 9, 309-315.
- Engel, J. F., Kegerreis, R. J. & Blackwell, R. D. Word of mouth communication by the innovator. Journal of Marketing, 1969, 33, 15-19.
- Frank, R. E., Massy, W. F. & Morrison, D. G. The determinants of innovative behavior with respect to a branded, frequently purchased food product. In G. Smith (Ed.), Reflections on progress in marketing. Chicago: American Marketing Association, 1964.
- Jackson, D. N. The dynamics of structured personality tests. Psychological Review, 1971, 78, 229-248.
- Jackson, D. N. Manual for the personality research form. Research bulletin 43. London, Ontario: University of Western Ontario, 1967.
- Jacoby, J. Personality and innovation proneness. Journal of Marketing Research, 1971, 8, 244-7.
- Kassarjian, H. H. Personality and consumer behavior: A review. Journal of Marketing Research, 1971, 8, 409-418.
- King, C. W. The innovator in the fashion adoption process. In G. Smith (Ed.), Reflections on progress in marketing. Chicago: American Marketing Association, 1964.
- King, C. W. & Sproles, G. B. The explanatory efficacy of selected types of consumer profile variables in fashion change agent identification. Unpublished working paper, Krannert Graduate School of Industrial Administration, Purdue University, 1973.
- Mischel, W. Personality and assessment. New York: John Wiley and Sons, Inc., 1968.
- Nunnally, J. C. Psychometric theory. New York: McGraw-Hill, Inc., 1967.

- Ostlund, L. C. Identifying early buyers. Journal of Advertising Research, 1972, 12, 25-30.
- Pessemier, E. A., Burger, P. C. & Tigert, D. J. Can new product buyers be identified? Journal of Marketing Research, 1967, 4, 349-55.
- Robertson, T. S. Determinants of diffusion behavior. American Marketing Association, Proceedings, 1967.
- Robertson, T. S. Innovative behavior and communication. New York: Holt, Rinehart and Winston, Inc., 1971.
- Robertson, T. S. The process of innovation and the diffusion of innovation. Journal of Marketing, 1967, 31, 14-19.
- Robertson, T. S., & Kennedy, J. N. Prediction of consumer innovators: Application of multiple discriminant analysis. Journal of Marketing Research, 1968, 5, 64-9.
- Rogers, E. M. Diffusion of innovations. New York: The Free Press, 1962.
- Rogers, E. M. & Shoemaker, F. F. Communication of innovations. New York: The Free Press, 1971.
- Rogers, E. M. & Stanfield, J. D. Adoption and diffusion of new products: emerging generalizations and hypotheses. In F. M. Bass, E. A. Pessemier, and C. W. King (Eds.), Applications of the sciences in marketing management. New York: John Wiley and Sons, Inc., 1968.
- Thorndike, R. L. & Hagen, E. Measurement and evaluation in psychology and education. New York: John Wiley and Sons, Inc., 1961.
- Worthing, P. M., Venkatesan, M. & Smith, S. Personality and product use revisited: An exploration with the personality research form. Journal of Applied Psychology, 1973, 57, 179-183.

AN EMPIRICAL TEST OF A SCALE FOR INNOVATIVENESS<sup>1</sup>

C. Samuel Craig<sup>2</sup>  
Cornell University  
and James L. Ginter<sup>3</sup>  
The Ohio State University

A scale developed to measure the trait of innovativeness was tested for predictive validity. The sample population of innovators were individuals who had purchased a new Mustang II early in the model year. Non-innovators were those individuals who purchased a new car other than Mustang II during the same time period. Results indicate that certain components of the scale (determined by factor analysis) did discriminate between sample populations.

### Introduction

Research into the temporal (see Rogers, 1962 and Robertson, 1971) and spatial (see Brown, 1968) aspects of the diffusion of innovation has taken many different tacks. Among the most intriguing studies, from the standpoint of consumer behavior, are those that attempt to identify characteristics of innovator populations versus non-innovator populations. Research delving into individual differences in the marketing literature has tended to focus on demographic characteristics, social interactions, and to a lesser extent, personality factors (Robertson, 1971, pp.100-101). Typically, personality factors account for a small and frequently non-significant portion of the variance in consumers' behavior. In a comprehensive review, Kassarian (1971) suggests that one of the major reasons for the lackluster performance of personality is the predilection for wholesale appropriation of pre-existing personality inventories that were not designed nor intended for marketing use.

The purpose of this study is to test the predictive validity of an innovativeness scale developed specifically to measure that trait. The actual development of the scale is covered elsewhere (Leavitt and Walton, 1974) and will not be dealt with here. The focus of this paper is to examine the usefulness of the scale in discriminating between a sample of innovators versus a sample of non-innovators.

### The Data

Potential respondents' names were gathered from lists of people who had purchased 1974 model cars from September, 1973, to December, 1973.<sup>4</sup> During January, 1974, individuals were first called by telephone and asked if they were willing to participate in a study about automobiles. Those who were willing to participate were then sent a six-page questionnaire which contained the innovativeness scale along with a number of other questions. Data gathered on the questionnaire also included: automobile A-I-O's (activities, interests and opinions), product use information, and demographic information. The questionnaire was mailed out to a total of 800 new car buyers. Of the 420 returned questionnaires 324 had fully completed innovativeness scales. Only these were included in the analyses.



Rather than consider a multitude of consumer characteristics, analysis focused on the usefulness of the innovativeness scale in identifying innovators. Innovators were operationally defined as those individuals who had purchased a 1974 Mustang II from its introduction in September, 1973, until December, 1973. In terms of Robertson's (1971) continuum of innovation the Mustang II is considered to be dynamically continuous. It does not merely represent a model change but, rather, a departure from what the Mustang had evolved to since 1964. It is not, however, a radical departure from existing automobiles such as an electric or steam-powered car would represent.

An attempt was made to have both samples identical except for the indulgence in innovative behavior. The alternative approach of selecting a random sample of area residents to compare against innovators was rejected. It was felt that use of a random sample would create a situation where the alternative explanation, i.e., that new car buyers, in general, are different, would be tenable. By considering two samples, both of whom purchased new cars, any differences are more likely to be a function of unique characteristics of innovators.

### Results

The results section is divided into three parts. Initially, a factor analysis (using varimax rotation) of the innovativeness scale is presented to determine its components. Then, the raw scale values are examined first as summed innovativeness scores and then collectively as predictor variables in a multiple discriminant analysis. Finally, respondents' factor scores were used in place of the individual scale items as input into several multiple discriminate analyses.

#### Innovativeness Factors

The actual factors are of considerable interest as they provide some indication of the dimensions contained by this scale of innovativeness (see Table 1). Factor 1, "New is Wasteful," seems to reflect a general distrust of anything new and different. In a sense it is a tradition-oriented factor and is indicative of a general resistance to new things. Perhaps the most important aspect of this factor is the notion that new things are inherently wasteful.

Questions loading on Factor 2, "Social Desirability," all indicate an extreme tolerance for others. Individuals who would agree to all the questions would tend to be very acquiescent. These five questions were taken from Crowne and Marlowe (1964) to detect those individuals who tend to give socially desirable responses to questions. As would be expected they all loaded together. Factor 3, "Novelty Seeking," indicates a general affinity for new things and new ideas. This factor would seem to be an important component of innovativeness, particularly when the product or idea is demonstrably new. Factor 4, "Risk Aversion," ought to be negatively associated with innovativeness. Individuals who are risk avoiders are not apt to be the first to try new products.

Factor 5, "Style Consciousness," is a component of innovativeness that seems to deal with people's awareness and concern about changing styles. Factor 6, "Satisfaction with Status Quo," on the other hand, suggests a desire to maintain the status quo. One puzzling aspect of this factor is the positive loading of, "When I see a new brand on the shelf, I often buy it

TABLE 1  
Factor Analysis of Innovativeness Scale<sup>5</sup>

Loading	Factor
Factor 1 (New is Wasteful)	
.62	The unusual gift is often a waste of money.
-.51	Some modern art is stimulating.
.51	I would rather not waste my time with some new idea.
.49	Buying a new product that has not yet been proven is usually a waste of time and money.
.49	I would like a job that doesn't require me to keep learning new tasks.
.45	The changing styles especially in clothes are a waste of money.
Factor 2 (Social Desirability)	
.66	I am always courteous even to people who are disagreeable.
.66	I have never been irked when people expressed ideas very different from my own.
.64	No matter who I am talking to, I am always a good listener.
.64	I am always willing to admit it when I make a mistake.
.55	I have never felt that I was punished without cause.
Factor 3 (Novelty Seeking)	
.74	I like to experiment with new ways of doing things.
.66	I like to fool around with new ideas even if they turn out to be waste of time.
.65	I like to try new and different things.
.45	When I see a new brand on the shelf, I often buy it just to see what it's like.
Factor 4 (Risk Aversion)	
-.72	I like to take a chance.
.68	When it comes to taking chances, I'd rather be safe than sorry.
-.65	I like people who are a little shocking.
Factor 5 (Style Consciousness)	
.80	I enjoy looking at new styles as soon as they come out.
-.64	The changing styles especially in clothes are a waste of money.
Factor 6 (Satisfaction with Status Quo)	
.72	I believe in leaving well enough alone.
.59	If people would quit wasting their time experimenting we would get more accomplished.
.47	When I see a new brand on the shelf, I often buy it just to see what it's like.

TABLE 1 (Continued)

Loading	Factor
Factor 7 (Other-Directedness)	
.78	If I got an idea I would give a lot of weight to what others think.
.64	I like to see what my friends and neighbors think of a product before I try it.

just to see what it's like." It could indicate that there is a brand-specific component and also a component dealing with society (or less specific factors) in general. Finally, Factor 7, "Other-Directedness," deals with looking to others for advice about products. It also seems to reflect some uncertainty about the soundness of one's own ideas.

#### Innovativeness Scale

The innovativeness scale was used to compare the respondents who had purchased a Mustang II with those who had purchased different cars during the same time period. The questions were rescaled so that all "more innovative" responses were in the same direction for all questions. Responses to the questions were first summed and Mustang II owners' totals were compared with those of Pinto, Vega, Torino, and Ford owners. There was no significant difference between the two groups.

In order to utilize the information collected by the scale more fully, discriminant analysis was used to compare the two groups on the 25 questions. This analysis did not consider the structure of the scale identified through the previous factor analysis, but it did take into account the multivariate relationship of the responses to the 25 original questions. The results showed that 3 of the 25 items were significantly different for the two groups at the .05 level. Mustang II owners agreed less with the statements, "I am always willing to admit it when I make a mistake" and "I like to fool around with new ideas even if they turn out to be a waste of time." Note that these questions loaded most highly on factors 2 and 3. Mustang II owners agreed more with "I enjoy looking at new styles as soon as they come out." This question loaded most highly on factor 5.

#### Factor Components Discriminant Analyses

The rather weak performance of the summed scores and the 25 individual items in the preceding comparative analyses prompted the use of the seven previously identified factors. If the factors actually represent constructs within the scale, one would expect greater differences between the groups. Using the respondents' factor scores on the seven factors, a discriminant analysis was run comparing Mustang II purchasers with all the other new car purchasers. Three of the factors were significantly different for the two groups (see Table 2). The Mustang II owners were lower on factors 2 and 4 (Social Desirability and Risk Aversion) and higher on factor 6 (Satisfaction with Status Quo). This improved discriminatory performance of the scale seems to indicate that the constructs identified through the factor analysis were operative in the purchase of the newly introduced brand.

The hypothesis tested in this analysis was that the two groups are equal on the set of factor scores. Although the test of significance is the appropriate test of this hypothesis, it is also interesting to consider the predictive power. Table 3 shows that the group of Mustang II owners was predicted with 62% accuracy and the alternative group was predicted with 60% accuracy. It should be noted that these figures were developed on the set of data on which the discriminant function was developed rather than a hold-out sample.

TABLE 2  
Discriminant Analysis:  
Mustang II versus Pinto, Vega, Torino, and Ford Using Factor Scores

Factor	Group Means		Partial $F^a$
	Mustang II (n=85)	Pinto, Vega, Torino, Ford (n=239)	
2 (Social Desirability)	-.237	.087	7.99
4 (Risk Aversion)	-.190	.071	5.06
6 (Satisfaction with Status Quo)	.193	-.032	4.09

<sup>a</sup>  $F_{.05, 1, \infty} = 3.84$

TABLE 3  
Prediction Results  
Mustang II versus Pinto, Vega, Torino, and Ford Using Factor Scores

Actual	Predicted		% Correct
	Mustang II	Pinto, Vega, Torino, Ford	
Mustang II	53	32	62
Pinto, Vega, Torino, Ford	96	143	60

The possible confounding effect due to aggregation of purchasers of very different types of cars in the previous analysis was eliminated by comparing owners of Mustang II with those of Pinto and Vega, some of its primary competitors. The results as shown in Table 4 are very similar to those of the previous analysis.

TABLE 4  
Discriminant Analysis:  
Mustang II versus Pinto and Vega Using Factor Scores

Factor	Group Means		Partial F <sup>a</sup>
	Mustang II (n=85)	Pinto, Vega (n=189)	
2 (Social Desirability)	-.237	.084	7.06
6 (Satisfaction with Status Quo)	.193	-.091	5.39
4 (Risk Aversion)	-.190	.070	4.57

<sup>a</sup>F.05 1,∞ = 3.84

A final comparison was conducted among the owners of brands manufactured by the Ford Motor Company. Mustang II owners were compared with those of the larger sized brands (Torino and Ford). The results (see Table 5) show that Mustang II owners were again lower on factor 2 (Social Desirability) and they were also higher on factor 5 (Style Consciousness).

TABLE 5  
Discriminant Analysis:  
Mustang II versus Torino and Ford Using Factor Scores

Factor	Group Means		Partial F <sup>a</sup>
	Mustang II (n=85)	Torino, Ford (n=50)	
2 (Social Desirability)	-.237	.099	4.51
5 (Style Consciousness)	.198	-.139	4.13

<sup>a</sup>F.05 1,∞ = 3.84

### Discussion

The summary of these results is interesting. In all of the comparative analyses, the Mustang II owners were lower on Social Desirability. The nature of these questions seems to indicate that these respondents tended to respond according to their feelings rather than responding in what might be deemed a more socially desirable way. They may also make decisions with less regard to what is expected of them and therefore accept more quickly (or seek out) a newly introduced brand. They also showed less risk aversion than owners of Pinto or Vega and Pinto, Vega, Torino, or Ford. This result is supportive in that those who were less averse to risk purchased the new brand. However, this factor was not significantly different between purchasers of Mustang II and Torino or Ford. This may be explained in part, since the purchasers of these three brands offered by the same company may have perceived little inter-brand

risk. An additional factor which was significantly different between these two groups was Style Consciousness, with the Mustang II owners scoring higher. This result is consistent with Ford's marketing objectives in offering a sporty, luxurious, and stylish economy car.

This application of the scale is also interesting in that in this instance, the factors associated with the wastefulness of new things, novelty seeking, and other-directedness were not significantly different for purchasers of the Mustang II. The Mustang II may not have been considered sufficiently new and different to be perceived as wasteful. Further, it may not have been perceived as sufficiently novel to bring out the novelty-seeking component. Also, in light of the earlier version of the Mustang, people may not have felt the judgment of others was essential (or non-essential) for the decision process. Given the magnitude of a new car purchase, an alternative explanation may be that other factors outweighed those related to the newness of the brand in the decision process.

### Conclusion

Certain components of the innovativeness scale appear to discriminate between innovators and non-innovators. In another situation with a different type of innovation, the other components might be significant. The results of the analysis suggest three conclusions. First, the innovativeness scale constructed by Leavitt and Walton has some predictive validity. The fact that a substantial proportion of the constructs identified within the scale were significantly different for Mustang II owners lends support for the scale and warrants further research on its application to other purchase situations. Second, the trait of innovativeness is not a homogeneous construct, but a conglomerate of a number of constructs. Further research, using different product categories, ought to be directed at determining which components of innovativeness are operating in different situations. Once this is established, marketers will have a firmer basis for identifying potential innovators and designing marketing strategies aimed at facilitating adoption of new products. Finally, the significant differences obtained strongly suggest the utility of developing scales aimed at measuring phenomena of interest to marketers. This would appear a far more fecund approach than attempting to apply less suitable extant scales.

### FOOTNOTES

1. This research was funded by the Division of Research, College of Administrative Science, The Ohio State University.
2. C. Samuel Craig is an Assistant Professor of Marketing at Cornell University.
3. James L. Ginter is Assistant Professor of Marketing, The Ohio State University.
4. Only individuals who had purchased a Mustang II, Pinto, Vega, Torino, or Ford were selected.
5. The seven factors explained 52 percent of the variance.  $n = 353$ .

## REFERENCES

- Brown, L. A. Diffusion dynamics. Lund, Sweden: Lund Studies in Geography, 1968.
- Crowne, D. P. & Marlowe, D. The approval motive: Studies in evaluative dependence. New York: Wiley, 1964.
- Kassarjian, H. H. Personality and consumer behavior: A review. Journal of Marketing Research, 1971, 8, 409-418.
- Leavitt, C. L. & Walton, J. Development of a scale for innovativeness. Working paper, The Ohio State University, 1974.
- Robertson, T. S. Innovative behavior and communication. New York: Holt, Rinehart, and Winston, 1971.
- Rogers, E. M. & Shoemaker, F. F. Communication of innovations. New York: The Free Press, 1971.

AN INSTRUMENT TO MEASURE CONSUMER SELF-ACTUALIZATION<sup>1</sup>

George Brooker<sup>2</sup>  
University of Wisconsin-Madison

Development of a personality measure based on Maslow's concept of self-actualization and designed for use in consumer research studies is described and the test is presented. Reliability and validity tests for the measure are described. Results of the tests are presented to enable researchers to evaluate the potential of the measure as a research tool for their own work.

Personality has been a subject of interest to marketers for some time. In attempting to understand and explain consumer behavior, marketers frequently have turned to personality as a variable having great potential importance for these tasks.

Recent personality-oriented consumer studies most often have used general clinical personality inventories or activity-interest-opinion ("life-style") measures to assess the influence of personality on the relevant behavior. The former approach has been criticized repeatedly, with the criticisms often followed by a call to marketers to develop their own personality measures (e.g., Kassarian, 1971, pp. 415-416). The use of life-style measures (which have been developed specifically for market-related studies) has helped add dimension and life to the paler demographic buyer-group differences often used in segmentation strategies. The major shortcoming of life-style research would appear to be the inadequacy of the present applications of life-style measures to advance or test theoretical personality formulations. This is not an unusual criticism of a highly practical personality research tool.<sup>3</sup> The development of a theory of life-style would seem to be a desirable goal to guide future research of this type (Frank, Massy, and Wind, 1972, p. 61).

One alternative to the measurement strategies presented above would be to use an instrument specifically developed for market-related studies which also represents a theoretical personality formulation. Such an instrument would be useful for testing a theory of personality in an appropriate market situation. For example, the Horney interpersonal orientation paradigm (as operationalized in Cohen's CAD Scale [Cohen, 1967]) would seem to be a useful theory for research into social influence in the consumer decision making process.

Of the numerous personality theories which have been advanced, one which has attracted a great deal of attention and interest is that of Abraham Maslow (Maslow, 1970). His conceptualization of the hierarchy of needs and the self-actualizing individual seem especially relevant as societies develop past the point where the material needs of the population are satisfied and other needs become more prominent.<sup>4</sup> It is Maslow's description of the self-actualizing personality which is used to develop the test described in this paper. This test represents a first attempt at development of a measure of consumer self-actualization.



## Review Of Maslow's Theory

The self-actualizing personality type combines several characteristics or traits. The self-actualizing person described by Maslow is a unique amalgam of hedonist and altruist, feeler and actor, detached and involved individual, judge of human nature and acceptor of human frailty. As embodied in a pure type, self-actualization, "...may be loosely described as the full use and exploitation of talents, capacities, potentialities, etc. Such people seem to be fulfilling themselves and to be doing the best they are capable of doing..." (Maslow, 1970, p. 150).

Maslow's theory is based upon a hierarchy of needs in which physiological needs are at the base and self-actualization needs are at the upper boundary. These needs, in order from lowest to highest level, are:

- Physiological needs
- Safety needs
- Belongingness and love needs
- Esteem needs
- Needs for self-actualization

According to the theory, satisfaction of lower level needs permits the needs at the higher levels to emerge, while unsatisfied lower level needs prevent the emergence of the higher level needs. The needs exist in a "hierarchy of prepotency," with the lower level needs being prepotent over the higher level needs. The needs which dominate (motivate) an individual at any point in time depend on the level of satisfaction attained for the other (lower level) needs.

As more of an individual's lower level needs are satisfied, they become less important in motivating behavior, and the individual becomes freer to seek satisfaction of the higher level needs. In Maslow's view, the freer the person is from lower level need motivation, the more closely that person approaches full psychological health and self-actualization. Sixteen character traits of self-actualizers have been identified by Maslow. The traits are presented, along with brief examples of how they might be operationalized, in Figure One. Further examples may be found in Maslow's text (Maslow, 1970).

It seems logical to assume that the more traits of a self-actualizer one possesses, the closer one comes to being self-actualizing for, with a single exception, no one trait is believed more or less important than another. The sole exception is the "peak experience" which is not necessary for psychological health; the absence of peak experiences does not prevent the individual from being self-actualizing. In fact, Maslow suggests the only difference between "peakers" and "nonpeakers" may be reflected in the life work they adopt (Maslow, 1970, p. 165). The other traits, in combination, sum to the potential of full psychological health or self-actualization.

### Other Measures Of Self-Actualization

Previous attempts to measure self-actualization or concepts based on theories related to self-actualization have taken several forms. In the study of consumer behavior, one such attempt at personality measurement utilizes the Q-sort technique (for an example, see Hamm and Cundiff, 1968). In theory, the technique establishes the psychological adjustment of the individual by comparison of a self-sort with an ideal-sort. Use of the Q-sort in consumer behavior has been criticized as being a measure of material well-being rather than self-actualization (Greeno, Sommers, and Kernan, 1973). The definition of self-

Figure 1. Description of Self-Actualization.

Characteristics	Sample Operational Descriptors
Efficient perceptions of reality and comfortable relations with it	Generally good ability to make correct judgments
Acceptance of self, others, and nature	Not particularly concerned over discrepancies between self and ideal self
Spontaneity, simplicity, naturalness	Simple, natural behavior
Problem-centering	Have mission in life which is often unselfish
Detached; need for privacy	Calm; serene; undisturbed by things which upset others
Autonomy	Relatively independent of physical and social environment
Freshness of appreciation	Maintain sense of wonder or awe for basic things of life
Feeling for mankind	Genuine desire to help the human race
Interpersonal relations	Deeper relationships with others than most adults
Democratic	Have friends in all walks of life
Discrimination between means and ends	Behave as though means and ends are distinguishable
Philosophical, unhostile sense of humor	Thoughtful, unplanned humor
Creativeness	Maintain a "naive approach" to perception
Resistance to enculturation	Maintain detachment from culture in which they are immersed
Resolution of dichotomies	Many oppositions seem to be merged to form unities
Peak experience	Feeling of limitless horizons opening to the vision

Developed from: A. H. Maslow, Motivation and Personality, 2nd, New York: Harper and Row, Publishers, 1970, Chapter 11.

actualization derived from the Q-sort methodology is not obviously related to Maslow's theory (Kassarjian, 1971, p. 413).

There have been no published reports of consumer behavior studies based on the need-level hierarchy. Attempts to measure need levels have been reported in studies of work situations, although there is some question regarding the ability of at least one of the measuring instruments to differentiate among Maslow's need levels (Roberts, Walter, and Miles, 1970; Payne, 1970; Lawler and Suttle, 1972). Scaling devices (Porter, 1961; Beer, 1966; Walsh, 1968; Goodman, 1968; Alderfer, 1969) and depth interviews (Hall and Nougaim, 1968) have been used with varying degrees of success in measuring need levels in managers and workers. Depth interviews would seem particularly difficult to use in large-scale consumer studies. Some modification of one of the need-level scales could be tried to make it suitable for a consumer behavior study.

Measures of self-actualization have been attempted through interviews (Bonjean and Vance, 1968), sentence completions (McKinney, 1967), and the Personal Orientation Inventory (Shostrom, 1964). Of these three types of measures, the Personal Orientation Inventory (POI) has been most widely tested and validated. It is a forced-choice type of scale containing 150 items. Within the POI there are two major scales and ten subscales, all of which are viewed as being partial measures of self-actualization. The notion of self-actualization which is tested is a composite of the views of Maslow, Fromm, Horney, Rogers, Riesman, and others, rather than a test of the views of one theorist. The length of the test would seem to make it awkward to use in many consumer behavior studies because of the possibility of respondent fatigue or boredom. The only study known (to this writer) to have employed the POI in exploring consumer behavior utilized students as respondents (Pasnak, 1968).

This brief review indicates a short consumer-related measure of self-actualization for use in consumer research would add to the tools presently available to measure self-actualization. In the following sections, the initial development of such a measure is described and several criteria for analysis are presented.

#### Instrument Development

For the initial stage of test development, a set of more than 150 items in the form of pairs of statements was written to represent the descriptors Maslow used to picture self-actualizers. (Examples of some of these descriptors were presented in Figure One.) Each pair of statements contained a self-actualizing and a non-self-actualizing choice. Content of many of the items was directly related to consumer actions or feelings. Items were written for all traits except the "peak experience." The items were pretested for clarity, familiarity of wording, and the like. The self-actualizing direction of responses was confirmed by a panel of judges composed of four faculty members of the Psychology Department of Northwestern University.

A convenience sample (usable N = 319) responded to the items which were presented in a five-point, A vs. B format. (The sample was composed of 151 male and 168 female respondents. The mean age of the respondents was 30.6, with a range of 17-82 years.) Care was taken to prevent respondent fatigue becoming a possible source of bias by varying the order of item presentation. Response-set bias was partially controlled by varying the self-actualizing response between A and B items in the test.

Responses to the items were dichotomized into self-actualizing and non-self-actualizing dimensions. Items were selected for inclusion in the test based on classical item analysis, which calls for selection of those items having the highest item-test correlations. Because correlation of an item with a test containing that item would overestimate the relationship, the partial (corrected) correlation of the item-test total was used. An attempt was made to have a large number of traits of self-actualizers represented by items in the test which was developed. Since the items representing the traits have a face validity of independence, the possibility of having inflated item-total correlations caused by item homogeneity is reduced.

In the item selection process, two items were taken from those representing each trait to form a tentative measure. The inter-item correlation matrix and the (corrected) item-total correlations were examined to determine which items in the tentative measure should be replaced. Criteria for item replacement were failure to reach significant levels in the (corrected) item-total correlations and/or significant negative correlations with other items in the tentative measure. Items dropped from the measure were replaced by other items representing the same trait as long as others were available. Numerous combinations of items were tried before the final set of twenty items was selected. The test which was developed using this method is presented in Figure Two.

#### Test Analysis

The test contains twenty items of the A vs. B variety. Scoring is done by assigning a value of one to each item receiving a self-actualizing response and zero to non-self-actualizing responses. Scores are summed over all items to produce a measurement index. Scores may range from zero to twenty. Distribution-free or nonparametric methods are recommended for data analysis. Interval scaling of the test score differences is not assumed and would be exceedingly difficult to test.

For the sample drawn to develop the test, test scores ranged from two to nineteen, with a mode of ten and a median of 10.4. A Chi-square test was performed on the distribution of scores to see how closely the distribution approximated the normal curve. The test was not significant (Chi-square = 4.08, 5 d.f.,  $p > .50$ ) indicating the scores obtained were not different from a normal distribution.

#### Reliability

Reliability was examined initially with a group of thirty MBA students (twenty-nine male, one female) enrolled at Northwestern University. A test-retest reliability figure of .57 (Spearman correlation,  $p < .01$ ) was obtained with four weeks between test administrations. Because the conditions of test administration did not appear to be similar (the first was given during a regular class period, the second was given just prior to the final examination when responses might be expected to vary due to nervousness, etc.), reliability was re-examined at a later date. A more heterogeneous sample of twenty-four persons responded to the test with five weeks between the first and second administrations. (The test items for the first administration were "buried" among a number of similar items.) The Spearman correlation coefficient for reliability obtained (corrected for ties) with the second group was .67 ( $p < .01$ ).

Figure 2.

The Consumer Self-Actualization Test  
(CSAT)

	<u>Corrected Item-Total Correlation</u>	<u>Characteristics, Trait</u>
1. *A. Information from publications like Consumer Reports is quite valuable to me in deciding on costly purchases.	.134 <sup>a</sup>	Detached
B. I find publications of little use to me; I prefer personal recommendations for expensive purchases.		
2. *A. Usually, I am not upset when products fail to meet my expectations.	.122 <sup>a</sup>	Accepting of Self, Others
B. I am often upset when products fail to meet my expectations.		
3. *A. When one helps a friend make a purchase, it's enough to know the assistance was needed.	.251 <sup>b</sup>	Feeling for Man- kind
B. When one helps a friend make a purchase, I think it's only proper that one is thanked for the help.		
4. *A. All people are worthy of respect.	.161 <sup>b</sup>	Democratic Char- acter Structure
B. Some people are not worthy of respect.		
5. A. It is more fun to give a gift than to decide what to give.	.228 <sup>b</sup>	Resolution of Dichotomies
*B. Deciding what to give for a gift is as much fun as giving it.		
6. A. I usually feel more confident when I know I am dressed in the latest fashion.	.404 <sup>b</sup>	Spontaneity, Simplicity
*B. I am at ease regardless of how I am dressed.		
7. *A. When I am shopping for myself, I make my decisions without help from others.	.172 <sup>b</sup>	Detached
B. I often look to people I know for help when buying something for myself.		
8. A. I like advertisements I see often more than those I have seen only a few times.	.150 <sup>b</sup>	Freshness of Appreciation
*B. I find myself trying to avoid advertisements I have seen before.		
9. A. Money may not be everything, but it's got a big lead over whatever is second.	.372 <sup>b</sup>	Means-Ends Relationships
*B. It's true, money can't buy happiness.		
10. *A. The things I desire for the good life are often different from those chosen by others in my economic class.	.159 <sup>b</sup>	Autonomy; Independence
B. The things I would choose for the good life are similar to the choices of others in my economic class.		

Figure 2 (Continued)

		Corrected Item-Total Correlation	Characteristics, Trait
11.	A. I am often bored.	.255 <sup>b</sup>	Freshness of Appreciation
	*B. I am seldom bored.		
12.	A. I like jokes that are slightly off-color.	.280 <sup>b</sup>	Sense of Humor
	*B. I like jokes that make me think.		
13.	*A. Activities like charity work and community service attract me.	.374 <sup>b</sup>	Problem centered
	B. I am so busy doing day-to-day things I can't be bothered thinking about volunteer work.		
14.	A. I prefer comedians who imitate famous people.	.190 <sup>b</sup>	Sense of Humor
	*B. I prefer comedians who comment on the present time.		
15.	A. When I go on a buying spree, I often regret it later.	.174 <sup>b</sup>	Accepting of self, Others
	*B. I seldom regret a buying spree.		
16.	*A. Being fashionable or chic holds no interest for me.	.286 <sup>b</sup>	Resistance to Enculturation
	B. It is important to me to be fashionable or chic.		
17.	*A. When I am nervous or anxious, I usually try to avoid buying things.	.168 <sup>b</sup>	Perceptions of Reality
	B. When I am nervous or anxious, I often find buying something new helps me feel better.		
18.	A. It is usually wise to present yourself to people in such a way that they will like you.	.271 <sup>b</sup>	Interpersonal Relations
	*B. It is unwise to present anything but your true self to people.		
19.	A. I am reluctant to try new products until I can find out if they're good.	.167 <sup>b</sup>	Perceptions of Reality
	*B. I often try new things just because they look interesting or good.		
20.	*A. I could probably live under any economic system and be just as happy.	.198 <sup>b</sup>	Resistance to Enculturation
	B. I can imagine myself being happy living only in one kind of economic system.		

a - indicates significance beyond the .05 level

b - indicates significance beyond the .01 level

\* - indicates direction of self-actualizing response

## Validity

All personality measures should be subjected to tests which determine how useful the measures will be in helping researchers make judgments about sampled populations. The various validity tests to which the measures are subjected form a multidimensional "picture." It is from this series of portraits that the researcher forms an overall impression of the usefulness of a personality measure for the research purpose at hand. The validity types normally used to evaluate psychological measures are content validity, predictive validity, concurrent validity, and construct validity (American Psychological Association, 1954). Predictive and concurrent validity often are discussed as two forms of criterion-related validity.

### Content Validity

Content validity refers to the adequacy with which the domain of the subject matter, such as consumer self-actualization, is sampled. The question of what an appropriate representation is from the universe of a psychological concept has not been resolved, and no rules have been specified for achieving proper representation (Robinson, Rusk, and Head, 1968, pp. 9-10). Proper sampling of content remains a problem in the development of psychological measures.

It would be impossible to represent every facet of self-actualization completely in a brief test. However, capturing a large number of the self-actualizer's traits in a test should be a good beginning toward an adequate portrayal of the concept. The test presented in Figure Two contains items from the sets written to match fourteen of the self-actualizer's traits identified by Maslow. Only the traits of peak experience and creativeness are missing.

### Criterion-Related Validity

Concurrent and predictive validity testing require validating test scores against some criterion behavior. The difference in the two types of validity merely reflects the time at which the criterion measures are taken (American Psychological Association, 1954).

Concurrent validity. Concurrent validity is examined by relating test scores to a criterion measure obtained at the same time responses to the test items are obtained. Often, studies of concurrent validity are addressed to the test's ability to discriminate between identifiable groups, such as the type of program in which a student respondent is enrolled.<sup>5</sup> The concept of self-actualization does not seem to lend itself readily to any such criterion. Maslow has suggested self-actualizers may be found in all types of work, making it difficult to validate the test against identifiable student groups, work groups, or the like. Concurrent validity may be tested using other criteria, however. One such criterion often used in marketing studies is consumer purchase behavior (Angelmar, Zaltman, and Pinson, 1972).

Based upon Maslow's theory, it was expected that self-actualizing consumers would be likely to buy products which serve society's interests and would take actions which would benefit others. The criterion measures studied, therefore, were purchase of products designed to minimize potential ecological damage to the environment and blood donation. In a study of detergent buyers, it was found that buyers of phosphate-free detergents (when compared with those who used no ecologically beneficial products) were likely (Chi-square test,  $p < .004$ ) to be high scorers (with scores divided at the median) on the Consumer Self-Actualization Test (CSAT). This result was confirmed by reported behavior of two other respon-

dent groups. In the group of detergent buyers, CSAT scores differentiated between those reporting use of lead-free gasoline and those reporting use of no ecologically beneficial products (Chi-square test,  $p < .003$ ). In a study of blood donation, a slight tendency (Chi-square test,  $p < .11$ ) was found for those reporting regular donations to be the higher CSAT scorers when compared with non-donors capable of blood donation.

Additional evidence of concurrent validity is found by examining the distribution of scores achieved by the population which was used for test development. No independent measure of concurrent validity was taken on the population, but there are two indirect indicators which may be used: differences in scores by sex and by age. White (1970) found a tendency for females to achieve higher scores than males on the Personal Orientation Inventory. A similar result was observed on scores for the CSAT, with females scoring higher than males at significant levels even when controlling for age (Chi-square test,  $p < .05$  for both younger and older groups).

The second indicator is based on analysis of respondent scores by age group. It was Maslow's belief that, while young people might be growing toward psychological health, full self-actualization was possible only in older people (Maslow, 1970, p. 150). To test for age-score differences, the sample was dichotomized into those less than twenty-six years old, and those twenty-six and older. A Chi-square test was performed, and the result indicated older respondents tended to achieve higher scores than the younger respondents ( $p < .015$ ). This result was confirmed, controlling for sex, by the Mann-Whitney U-test ( $p < .05$ , one-tailed test, for both males and females).

Predictive validity. Tests of predictive validity compare results obtained when predictions made from test scores are evaluated against a subsequent behavioral measure. No tests of predictive validity have been performed on this personality measure.

#### Construct Validity

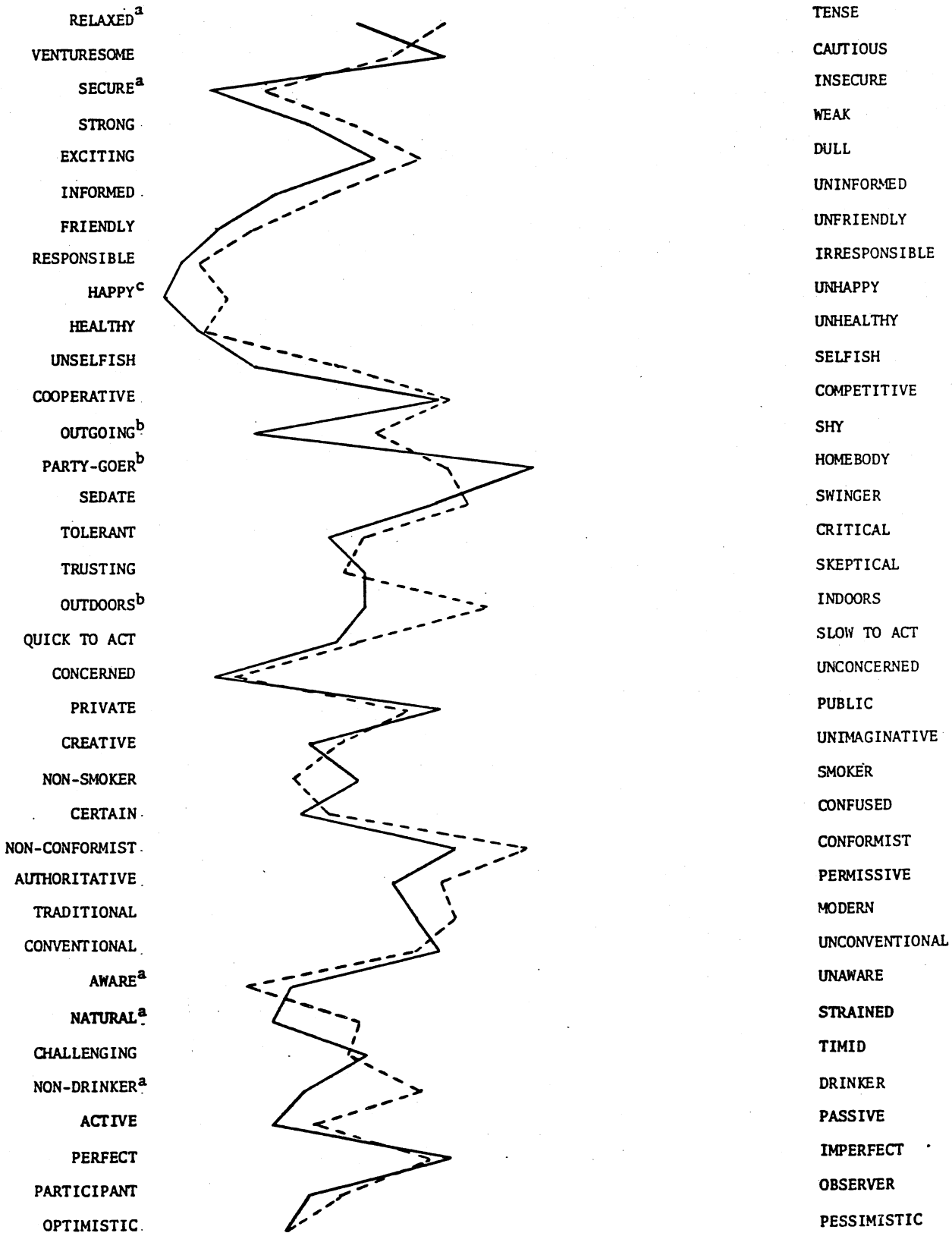
Construct validity tests the psychological measure and hypotheses based on the psychological construct at the same time. Such tests may correlate the measure being evaluated with other psychological measures, or the tests may be based on operationalization of the concept as it is expressed in characteristics measured in an independent manner. That is, construct validity explores the psychological qualities measured by a test as they are expressed through separate confirmation of the presence of characteristics one would expect in individuals possessing the qualities believed to be gauged by the psychological measure being tested. It is through the latter validation procedure that construct validity has been tested with the CSAT.

The CSAT and a set of semantic differential scales for self-description were given to a convenience sample of twenty-four persons. (The group of people was evenly divided between males and females. The range in ages was twenty-one to fifty-four, with a mean of 34.6 years.) The CSAT and the self-descriptive semantic differential measure were separated by two other semantic differential scales in the test booklet. Scores on the CSAT were divided at the median, and t-tests were calculated for differences in means on the semantic differential items. A graph depicting the differences in means obtained is presented in Figure Three. (Several items have been reversed for the figure from the test procedure to improve visual interpretability.) Of the thirty-six semantic dif-



Figure Three

SELF DESCRIPTIONS OF HIGH AND LOW SCORERS ON CSAT  
 (KEY: HIGH SCORERS ————  
 LOW SCORERS - - - - -)



a:  $p < .10$  (one tail test)  
 b:  $p < .05$   
 c:  $p < .025$

ferential items, nine showed significant differences (at the .10 level) between high and low scorers on the CSAT. The picture presented by the high scorers on eight items (more relaxed, more outdoorsy, more secure, happier, more outgoing, more homebodyish, more natural, non-drinkers) seems to parallel Maslow's view of self-actualizers being generally psychologically healthy. It would not take a great stretch of the imagination to interpret the one apparent reversal in the significant differences (more unaware) as psychologically health, as well.

#### Conclusion

The reliability and validity studies reported here present several criteria by which researchers may evaluate the potential research value of the CSAT. It should be recalled that this is a first attempt at development of this personality measure, but even fairly crude devices can be useful in empirical research. Research by others using the CSAT would be helpful in providing further information on the measure's validity and directions for its improvement.

#### FOOTNOTES

1. The author gratefully acknowledges the Richard D. Irwin Foundation and the Vogelback Computer Center of Northwestern University for financial support of much of the research and analysis presented in this paper.
2. George Brooker is Assistant Professor, Graduate School of Business, University of Wisconsin-Madison.
3. The problem of finding a test with both practical value and theoretical significance has been noted by Steiner (1966, p. 208).
4. Lazer (1969) appears to have been referring to this theory of personality when he suggested a society which has had its material needs satisfied might then be free to seek higher need satisfactions. The increasing attention given to the problems of advancing societal causes probably is indicative of the feeling of many scholars that the age of affluence may soon be superseded by an era of empathy. See, for example: Kotler and Levy (1969); Kotler and Zaltman (1971); and Peters (1973).
5. For example, Cohen (1967) tested the CAD Scale on students enrolled in social welfare courses, business administration programs, and geology programs, finding students in each of these programs scored higher on the expected portion of the scale than did their peers in the other programs. Similarly, the social character scale for inner-other directedness (Kassarjian, 1962) was tested on student populations for concurrent validity.

#### REFERENCES

- Alderfer, C. P. An empirical test of a new theory of human needs. Organizational Behavior and Human Performance, 1969, 4, 142-175.
- American Psychological Association. Technical recommendations for psychological tests and diagnostic techniques. Psychological Bulletin Supplement, 1954, 51, 1-30.

- Angelmar, R., Zaltman, G., and Pinson, C. An examination of concept validity. Proceedings, Third Annual Conference, Association for Consumer Research, 1972.
- Baer, M. Leadership, employee needs and motivation. Bureau of Business Research Monograph No. 129. Columbus, Ohio: The Ohio State University, 1966.
- Bonjean, C. M., and Vance, G. G. A short-form measure of self-actualization. Journal of Applied Behavioral Science, 1968, 4, 299-312.
- Cohen, J. B. An interpersonal orientation to the study of consumer behavior. Journal of Marketing Research, 1967, 4, 270-278.
- Frank, R. E., Massy, W. F., and Wind, Y. Market segmentation. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1972.
- Goodman, R. A. On the operationality of the Maslow need hierarchy. British Journal of Industrial Relations, 1968, 6, 51-57.
- Greeno, D. W., Sommers, M. S., and Kernan, J. B. Personality and implicit behavior patterns. Journal of Marketing Research, 1973, 10, 63-69.
- Hall, D. T., and Nougaim, K. E. An examination of Maslow's need hierarchy in an organizational setting. Organizational Behavior and Human Performance, 1968, 3, 12-35.
- Hamm, B. C., and Cundiff, E. W. Self-actualization and product perception. Journal of Marketing Research, 1969, 6, 470-472.
- Kassarjian, H. H. Personality and consumer behavior: A review. Journal of Marketing Research, 1971, 8, 409-418.
- Kassarjian, W. M. A study of Riesman's theory of social character. Sociometry, 1962, 25, 213-230.
- Kotler, P., and Levy, S. J. Broadening the concept of marketing. Journal of Marketing, 1969, 33 (1), 10-15.
- Kotler, P., and Zaltman, G. Social marketing: An approach to planned social change. Journal of Marketing, 1971, 35 (3), 3-12.
- Lawler, E. E., and Suttle, J. L. A causal correlation test of the need hierarchy concept. Organizational Behavior and human performance, 1972, 7, 265-287.
- Lazer, W. Marketing's changing social relationships. Journal of Marketing, 1969, 33 (1), 3-9.
- Maslow, A. H. Motivation and personality. (2nd ed.) New York: Harper and Row, 1970.
- McKinney, R. The sentence completion blank in assessing student self-actualization. Personnel and Guidance Journal, 1967, 45, 709-713.
- Pasnak, M. F. D. Fashion innovators compared with non-innovators on clothing attitudes, self-actualization, and tolerance of ambiguity. Unpublished doctoral dissertation, The Pennsylvania State University, 1968.

- Payne, R. Factor analysis of a Maslow-type need satisfaction questionnaire. Personnel Psychology, 1970, 23, 251-268.
- Peters, W. H. Who cooperates in voluntary recycling efforts? In Thomas V. Greer (Ed.), Combined conference proceedings, American Marketing Association, Spring and Fall, 1973.
- Porter, L. W. A study of perceived need satisfactions in bottom and middle management jobs. Journal of Applied Psychology, 1961, 45, 1-10.
- Roberts, K. H., Walter, G. A., and Miles, R. E. A factor-analytic study of job satisfaction items designed to measure Maslow need categories. Proceedings, 78th Annual Convention, American Psychological Association, 1970.
- Robinson, J. G., Rusk, J.G., and Head, K. B. Measures of political attitudes. Ann Arbor: Survey Research Center, 1968.
- Shostrom, E. L. An inventory for the measurement of self-actualization. Educational and Psychological Measurement, 1964. 24, 207-218.
- Steiner, G. A. Consumer behavior: Where do we stand? A psychologist's appraisal. In J. W. Newman (Ed.), On knowing the consumer. New York: John Wiley and Sons, Inc., 1966.
- Walsh, T. M. Maslow's theory and motivation: A method for testing the constructs. Unpublished doctoral dissertation, University of Oregon, 1968.
- White, H. M. An investigation of some characteristics of high and low self-actualization and their relationship to alienation from self and society. Unpublished doctoral dissertation, University of California, Berkeley, 1970.



THE ADVERTISING PRETEST AS PART OF A MULTIMEASURE, MULTIMETHOD, MULTISITUATION  
VALIDATION AND APPLICATION RESEARCH SYSTEM

Michael L. Ray  
Stanford University

This invited paper was developed as part of a workshop on pretesting at the ACR meetings. The author's contention is that a collaboration of managerial, behavioral and quantitative people in advertising can be established with a research system that attacks the important "problems of the middle range" that occur repeatedly in advertising. The research system would include, in order: (1) examination and development of the problem structure, (2) analysis of available literature, (3) laboratory pretesting, (4) computer simulation application of pretesting results, (5) field experimental validation of lab and simulation predictions, and (6) campaign monitoring. The system includes managerial aspects with multimeasures representing managerial goals, behavioral aspects with multimethods representing the main contributions of the behavioral sciences, and quantitative concerns with multisituations bringing in the quantitative sciences. This approach is supported by an extensive research program with one dozen lab pretests which balance natural conditions and research needs, a media modeling study, two extensive field studies, and several monitoring projects.

The advertising pretest, when it is used correctly, is part of a managerial process. It can be used to deal with general, recurring questions that keep coming up in specific advertising decision situations. The pretest is quite inefficient if it is used in isolation or if a new form of pretest is developed for each individual situation. So the goal of our research at Stanford has been to develop a standard pretesting format for general managerial problems in advertising.

Thus far we have done about a dozen developmental studies involving about 3000 respondents. The general technique we have developed has been tentatively validated by two large-scale field experiments, one on print advertising and one on television advertising. Our data have been used in media model runs with positive results.

The purpose of this brief statement is to provide an overview of our research: our assumptions, main findings and implications for advertising pretesting. For those interested, there are three papers which give greater detail on the research (Ray, 1972, Ray et al. 1973, Ward and Ray, 1974).

#### Types of Pretesting

There are essentially three types of pretests which come at different stages of the advertising decision-making process. The first is what might be called developmental pretesting. This is done as the advertising is being created. It should be the province of the creative people within an advertis-

ing agency, because developmental pretesting is used to determine whether certain components of advertising work. Typically, the materials are unfinished advertising or advertising in a rough form. The samples are usually accidental. The measures tend to be either gross physiological ones or, on the other hand, consumer "expert" opinions about the content of advertising.

The second form of pretesting is what might be called selection-scheduling pretesting. This is done with finished advertising, and its goal is to determine which of several alternatives should be run and how those alternatives should be run or scheduled. Here the alternatives are in as finished form as possible, the sampling is more representative of the audience, and measurement is made to determine realistic communication response rather than physiological indicators or consumer opinions of advertising effectiveness. The name of the game here is to predict to the response that is in fact the objective of the campaign. Because of this goal, it is necessary to make the exposure conditions and measurement as natural as possible, considering the usual "laboratory" type of setting for the test.

The third type of pretesting might be called limited posttesting. The best example of this sort of approach is the day-after-recall test. This is usually used at a later stage in the decision-making process when the alternatives have been narrowed down to a very few, often times one. In a sense, then, this is used as a final check in many cases to assure decision makers that the commercial they are about to run is an excellent one. While some companies such as Proctor & Gamble use limited posttesting on a regular basis for selection and scheduling, this tends not to be the norm. The advantage of limited posttesting are natural exposure and quick measurement. The disadvantages are a limitation in both measurement and possible variation of alternatives. Also, sometimes, the cost can be too high.

The Stanford research has opted for the middle ground in pretesting; that is, we have worked with just the selection-scheduling type of pretest. Our feeling has been that it is desirable to have finished advertising to determine actual potential of ad response. The laboratory setting of the selection-scheduling pretest offers the potential for control of those variables which are critical in advertising effectiveness. Such pretesting can be done quickly with less cost than the limited posttesting type. It is possible to make the conditions of exposure quite natural and measure on a number of levels so the full texture of communication response can be gauged. Most important for us, the selection-scheduling type of pretest has the potential of being used in a research system involving not only pretests but also behavioral analysis, media models, field experimentation, and campaign monitoring. Such a system promises the potential of validation of both the pretest and the models. The end result of such an approach should be not only better pretesting but also more extensive, efficient utilization of both behavioral science hints and media models in advertising decision-making.

#### The Research System Context for Pretesting

The Stanford research started with the general goal of developing a pretesting technique that could be used to determine the repetition response function for advertising in specific situations. As this work began, we began to realize that the repetition function or wear out problem in advertising was really only one of a number of general recurring problems that needed to be attacked in a more systematic way than had previously been done.

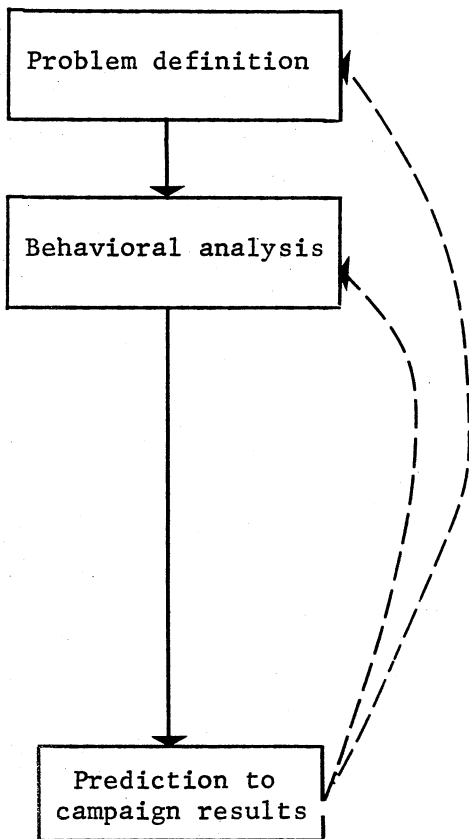
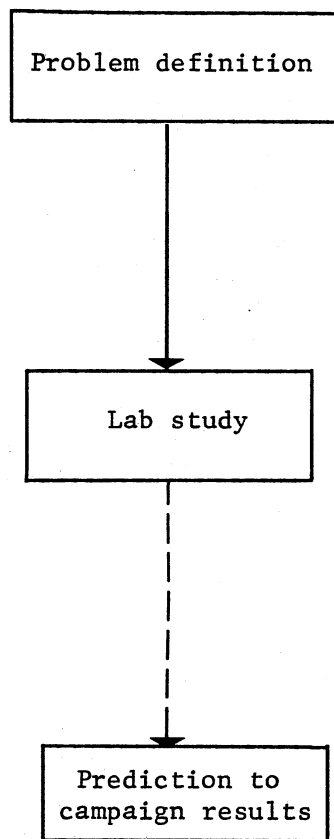
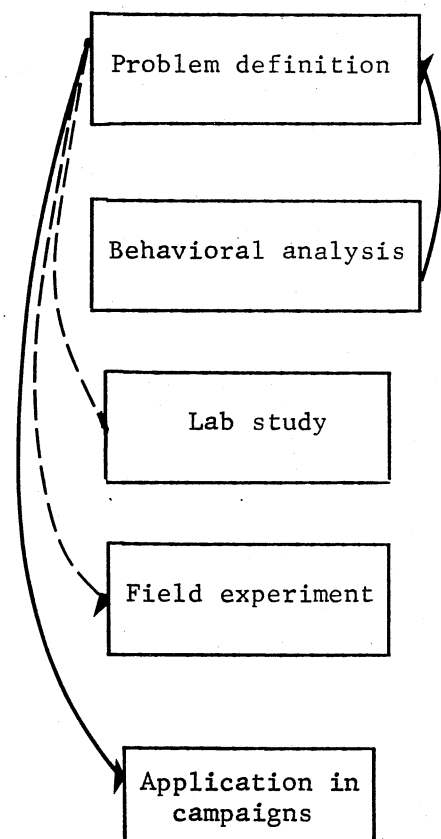
The situation as we saw it for these general recurring problems could be depicted as shown in the chart on the following page. In essence, there were three general types of attacks on advertising problems, none of which was totally satisfactory. The first was the advertising media modelers' or management scientists' approach. These people started with a well-defined problem, such as advertising scheduling, and then proceeded to develop a clear conceptualization of the problem area by working intensively with managers who understood it. On the basis of that discussion and references to simple-minded behavioral science ideas, media models were created. In some situations, these models are used to plan campaigns, but they are never satisfactorily validated.

The second approach shown on the following page is the one this paper and workshop are most concerned with. This is the typical copy testing approach. Unlike the management science modeling one, the copy testing approach does not specify a problem that can be dealt with in a continuing and ordered way. The usual question given to a copy tester is a very short term and narrow one such as: Is Ad A Better Than Ad B? This kind of question is almost immediately examined in a copy test without: (a) any intervening examination of past experience relative to the general types of ads involved or (b) behavioral science information which might provide hints as to whether either or both of the alternatives should be considered at all. The copy testing procedures that are used are also deficient in several ways. Usually copy tests are done with accidental samples, without the normal surrounding material that is involved when ads are actually exposed, without the variations of exposure that allow a scheduling decision (that is, there tends not to be any repetition variation or variations in program type or variation in media type or variation in competitive setting), without measurement on the several levels that determine advertising response in an effective way, and without delayed measurement or measurement of purchasing action. What is being described here is the "one shot immediate measurement copy test." Not only is such a copy test typically one shot in terms of exposure, but there is also a tendency for copy testing researchers to seek the single ideal measurement. This is a search which could be likened to the search for the Holy Grail. In fact, a single measure can measure only a single response, and no single measure, whether it be physiological response or response time or anything else, can predict directly to a field or campaign response under natural conditions.

In addition to having difficulties in terms of the problem statement, the lack of behavioral science and experience review, and the characteristics of the copy tests themselves--the copy testing approach suffers from a lack of actual application. The dashed line in the chart is meant to indicate that only sometimes are the results of copy tests actually applied to campaign decision making. The managerial situation in which copy tests are done often leads to their purpose being something other than improving the quality of the campaign. It seems that psychological, organizational and political factors are often a greater reason for copy tests than anything else. Since these copy tests do not have the variations that might allow them to be related to the field, it is very difficult for copy testers to say anything more than Ad A seems to be better than Ad B. This provides very little direction in terms of how the ads would operate in an actual schedule in an actual campaign. And because there is very little use of specific copy testing information in campaign development, there is also very little feedback from the campaign to subsequent decisions or to the copy testers themselves.

The third type of approach to advertising decision making problems is the academic application of behavioral science ideas. The main problems with this



MODEL BUILDINGCOPY TESTINGPROPOSITIONS APPLICATION

The state of behavioral application in advertising.

approach, as can be seen in the chart, is that the behavioral scientist tends to start with propositions and then attempt to find a problem to which to apply the propositions. Beyond that difficulty, the propositions sometimes get applied in actual campaigns, but seldom is there application in copy tests or field experimentation. And feedback is almost non-existent.

When one attempts to solve a particular problem by using the three approaches shown in the chart, he or she often finds that these approaches give very different answers to general problems. For instance, in the case of the repetition function problem, we found that each approach gave a different answer to the question of whether the repetition function was affected by different advertising situations. The management science model builder tended to use a single function, adapted from the findings of nonsense syllable verbal learning research, to depict all of the possible responses in all possible situations. So the model builders answer to the question was typically: "No, there are no differential effects on the repetition response function due to the situation." The copy tester, on the other hand, earns his living by finding differences between commercials. So his answer to the repetition function question, even

though he or she seldom does any copy testing with repetition, is "Yes, there are big differences in response to advertising." The academic behavioral science applier tends to see a great deal of complexity in the world. Since, as shown on the chart, there is really no research done in this approach, the behavioral scientist answering the question usually said: "Well maybe, there might be response function differences."

Given the ambivalent state of affairs with regard to the repetition function problem and others which recur in advertising, we felt that there was a need to develop a research system which could be applied to such problems. The basic assumption of this research system is that the problems can be most efficiently handled with a combination of the best aspects of each of the three approaches shown in the chart. One current version of our recommended research system is in Figure 1, which originally appeared in Ray (1972, p. 476). In this case, the management problem of concern was: "What kind of messages should be used in highly competitive situations?" This is representative of the type of general, recurring problem that should be studied by such a research system.

The second stage of the research system is one in which behavioral science and past experience is tapped to develop alternative strategies, in the case of Figure 1, this meant the development of alternate message strategies.

The third stage is the one involving copy testing. In Figure 1, it is called the laboratory experimentation stage. The function of copy testing in such a research system is to take the alternatives developed from the second stage and test the way they have been implemented in the specific situation of concern. Copy testing or laboratory experimentation is emphasized because it provides a quick and low cost method to separate out those alternatives that should be considered further. It also allows the decision maker to see how the ideas from the behavioral area and past experience apply in this specific setting. It is with these types of goals in mind that the pretesting research and techniques of our project have been developed.

The fourth stage of the research system has to do with media model runs. In the few cases in which copy testing results have been given a validation test in the field, there has been a naive attempt to see if the results in the laboratory as they stand apply directly to the results in the field as they stand. There are several problems with this sort of validation approach. The fourth stage of the research system is an attempt to eliminate some of these problems, which come from lack of attention to field situation variables in making the predictions to the field. Media models have been developed which include as part of their data base, many of the factors which might affect the operation of various commercials in the field. By using response function estimations from the lab within the context of the other media models inputs, it is possible to make very precise and realistic predictions to the field.

The fifth stage in the research system is field experimentation. It provides a solution to another of the validation problems of pretests. That is, pretests are often asked to predict to a situation in which the very variables which are manipulated in the lab are not cleanly manipulated. It is unreasonable to expect a variation in copy to show up in a test marketing situation in which these variations in copy have not been experimentally manipulated. The field experimentation recommended by the research system would not have to be done often, but it would have to be done from time to time to accurately validate both the copy tests and the media models. The

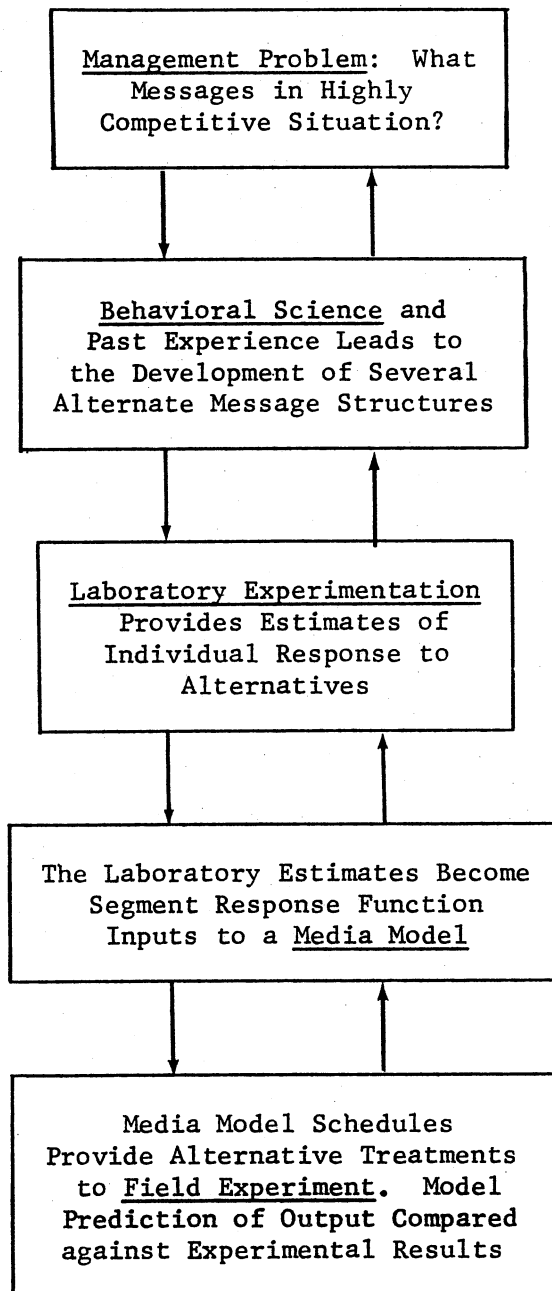


Figure 1. Outline of proposed research.

article from which this figure came indicates how such field experimentation might be developed to test these two components of the research system.

There is one more stage of the research system which is not shown in Figure 1. This is the campaign monitoring stage. Once the previous five stages of the research system are done, there would be a great deal of information on the particular alternatives that survive and are actually run in campaigns. Continuous monitoring systems, especially of the repetitive survey type, can provide additional information which might be utilized in the earlier stages of the decision process.

#### The Laboratory Experimental Technique

When pretesting is considered as part of a decision and research system of the sort partially depicted in Figure 1, there are certain requirements for it that otherwise would be ignored. There is a need to present materials in a very natural way with all of the key situational variables operating while at the same time there is clear experimental control and definitive measurement of response.

The procedure we have used in the Stanford program is an after-only experimental design with a double-blind cover story. Typically there are repetitive exposures of test commercials embedded in a stream of messages, and there is multiple measurement after test exposures to determine the nature of communication response.

This technique has been used for several problem areas and in a variety of types of locations. These have included mobile units in shopping centers, store fronts, schools, and central location research facilities. Respondents are recruited to participate in studies that are a cover for the actual intent of the research. In most of the research we have done the "Shopping of the Future" cover story has been used. When they arrive at the facility, respondents are given materials on possible alternative means of cable television and teletype-telephone home shopping approaches that might be used in the future. They are then told that they will see a demonstration of such an approach which consists of a stream of messages presented on a futuristic television screen display. In some studies we have departed from the "Shopping of the Future" cover. We have told respondents that the research was being conducted to investigate television violence and humor or television ratings. In the case of these studies, respondents then saw normal program material with our test commercials embedded in the programs.

Following whatever presentation the respondents see, they fill out a self-administered questionnaire that first has questions related to the cover story and then questions on play-back of the messages, cognitive response to the messages, attitude and purchase intention, and cued recall and response to the messages. Also included in the questionnaire are questions on respondent characteristics and past experience with the products and brands involved. Sometimes there are behavioral measures which follow the test itself.

Several aspects of this procedure should be underlined. It combines the advantages of experimental control with an effective cover story and relatively natural exposure conditions. The respondents are not told to concentrate on the commercials themselves. Rather they are concerned with the general presentation, which is closely connected to watching television normally. In addition, they usually view the commercials in a "living room" setting along with two or three other individuals. At the same time that this natural setting is achieved,

it is possible to show respondents commercials at 0 through 3 exposures in competition with the normal commercial fare that is offered on television. In fact, it is possible to add print ad versions of commercials and increase exposure up to 10. The studies that have been done with the technique have included competitive messages, multi-media effects, differences in cover story and degree of attention directed to messages, variation in distraction, and in the types of viewing groups.

### Review of Results with the Technique

By doing pretesting within the setting of the research system just described, it is possible to develop a more realistic assessment of the communication response that occurs in specific advertising situations. As the repetition project has developed, we have moved from a belief in the straight learning hierarchy-of-effects ideas with common exponential functions of response toward one in which there are a variety of hierarchy possibilities with unusual functional relationships. A course of that expanding realization has run through the studies that are outlined below. More detail on these studies is found in Ray (1972), Ray *et al.* (1973) and Ward and Ray (1974). The paragraphs below simply give a general indication of the development of our thinking.

Studies I through III: Repetition Pretest Technique Development. In the first three studies we were concerned with determining whether we could get repetition response functions in the laboratory setting that were realistic and consistent from study to study. The first study involved 18 different test advertisements, three ads for three brands in each of six product categories. Not only were there repetition effects but there were also interesting differences across product categories, ad types, life cycle stage of the product, and across specific ads. The second study was done with a subset of the advertisements from the first study, and the results were consistent with those that were developed in the first one. The third study concentrated on "refutational" versus "supportive" types of advertisements. There were pairs of ads for five different brands in five different product categories. In addition, there was a test of a color versus a black and white version of a campaign (four advertisements) for a grocery product. There were strong refutational versus supportive differences when considered by usage groups. Unlike the results in study one, which indicated that there was no consistent color versus black and white difference, this study found that the color campaign was more effective in generating gross awareness of the advertising, while the black and white campaign, if recalled, was recalled at a greater depth. The conclusion of these three studies was that we had a technique that worked quite well and gave us interesting and, on a face validity basis, reasonable results.

Field Study 1. This study was a replication with variation of the mail print advertisement field experiment that was done originally by H. Zielske in 1959. The study ran over a 13 week period with respondents receiving mailings on a weekly basis. Embedded into the mailings were weekly, bi-weekly and monthly schedules of advertising for six different ads or ad campaigns. All of the advertising used in this print field study was previously used in studies I through III. In addition to developing data supportive of Zielske's earlier study, Ed Strong, who did this study, was able to develop a scheduling simulator that could be used to evaluate various advertising schedules. His research is reported in an article in the November, 1974, Journal of Marketing Research. More important for the present purposes is the fact that his findings were also supportive of the results of studies I-III. One thing that was learned, however, is that strong response measures from the laboratory were likely to be good predictors of weak response measures in the field. For instance, in the color versus black and white test, the black and white ads that did well on depth

response in the lab tended to out-perform the color ads in ad and brand awareness in the field. This kind of "translation" from the lab to the field has been necessary in all the validation work we have done.

Model Application Study. Since one of the goals of the repetition project was to develop response function information for media models, it seemed reasonable that we should attempt to apply some of the laboratory findings to runs of the well-known media model MEDIAC. The data from study III on differences for refutational versus supportive response in purchase intention in various usage groups was applied in a number of ways to a series of runs of the model. The general finding (reported in detail in Part II of the December, 1971, Management Science) was that the model ran much more efficiently and produced better results when the more textured data from the laboratory pretests were used as opposed to a single function run of the model. This study gave some support for the attempt to develop more realistic data from advertising pretests than is usually obtained.

Studies IV through VI: Repetition with Variation. These three studies, which were conducted by Roger M. Heeler, were an extension of the use of the repetition laboratory technique in a new setting and for new purposes. These were the first studies that were not done in a mobile unit parked in a shopping center. Instead Heeler did these studies in a central research facility. In addition, these were the first studies done with television advertising. Heeler studied in one project repetition with variation in message, in another repetition with variation in media and in a third study the effect of repetition on perceptual maps. Thus this set of studies pushed the ability of the general technique to its boundaries. Heeler found that the effect of television advertising was more dramatic than the effect of magazine advertising. He also learned that there was usually a need for some print in a "campaign" in order to achieve maximum effect. But the mixed media campaigns did not work in all advertising situations. Heeler was able to develop a general simulation of effects, given his data on mixed media. The effects of variation in message were not as dramatic. It is not always true that variation in message can extend the life of a campaign. By varying the amount of competitive advertising, Heeler was able to show that response was quite different depending on the measure. More competitive advertising actually helped the recall of test advertising. But it had a negative effect on attitude and purchase intention measures.

Study VII Methods of Continuous Repetition Response Measurement. A small scale developmental study was done to determine the effects of asking for response during the exposure to repetitive advertising. In the standard design we normally use, measurement is done only after exposure to messages. There would be significant advantages if it were possible to get some measurement during exposure without affecting the exposure itself. Two methods were tried in a small pilot project done at the U.C. Berkeley management science laboratory. One of the measures was galvanic skin response. By attaching respondents to a GSR monitoring system, it was possible to see if resulting gross arousal changes had any pattern during the advertisement exposures. The other during-exposure measure was a teletype-activated scale by which respondents can indicate how interested they were in the material they were seeing. There were essentially four experimental treatments: GSR only, teletype only, both GSR and teletype, and normal after-exposure measurement only. The analysis of this study is not complete, but surprisingly enough, there does not seem to be a great effect caused by the during-exposure measurement. This promises to add dimensionality to future studies.

Study VIII Repetition of Political Advertising. Michael Rothschild did a study in a shopping center storefront in which the key variables were levels of political contest and political involvement of respondents. His was the first study in which we clearly observed two different types of hierarchy of response within the same experiment. For the presidential advertising, increased repetitions produced standard learning hierarchy results across the cognitive, affective and conative measures in the study. For the more low involvement state assembly race, there were effects on the cognitive and extremely strong effects on the attitudinal measure. This was similar to what would be predicted by Herbert Krugman's low involvement learning hypothesis. These results are reported in Ray *et al.* (1973) as well as in an article by Rothschild and Ray in the July 1974 issue of Communication Research. Rothschild's study also included products that were similar to the classifications in Studies I and II. Again, there was support for the findings in that initial study, as well as some indication of the comparative response to product advertising as opposed to political advertising. Presidential campaign advertising seems to operate in a more high involvement way than the product advertising; whereas the congressional and state assembly advertising seems to operate in a more low involvement manner than the product advertising.

Political Campaign Monitoring Studies. As part of his dissertation research, Michael Rothschild also collected extensive data on one political campaign for the Senate in a midwestern state and on a number of ballot propositions in California. Although these data were not in a form that allowed a precise test of the ideas that were supported in the laboratory experiment, there was some support as far as the analysis could go.

Media Vehicle Exposure Value Research Proposal. Professor Alvin J. Silk of MIT and I developed a proposal for a Marketing Science Institute project on determining the extra value or qualitative value of medical publications. The research proposed was quite similar to that used in the repetition area, and it gave further testimony to the flexibility of this particular copy testing technique.

Studies IX and X: Anti-drug Abuse Advertising Pretesting Procedure. Jerome B. Reed did two large-scale laboratory projects in which repetition was not a variable, but distraction, attention to commercials, competition, audience, and message type were variables. Contrary to previous research in this series, the advertising was placed in the context of program material, and the cover story consisted of the "television violence-humor project." Parents, junior high school and senior high school students were the three kinds of audiences for the study, and they came to central school and organization locations for the test interviewing. These two large-scale laboratory studies convinced us that the response to advertising is much more complex than would be indicated by even the three-orders model mentioned in Ray *et al.* (1973). By including cognitive response measures in various conditions of distraction and audience type, we were able to determine that some messages were able to be recognized and affected later behavioral responses, while at the same time generating quite negative cognitive responses and little attitude change. Other messages created a generally positive response but nothing that seemed to be indicative of a behavioral one. These findings, which are presented in more detail in Ward and Ray (1974), indicated a planning procedure in which the three-orders hierarchy model is used as an initial planning step but copy testing with multiple measurement is essential to determine the nature of communication response in each specific advertising situation. This research argued very clearly for the textured sort of copy test and against the single measure, single exposure, unnatural setting type of copy test.

Field Study II: Split Cable Experiment with Anti-Drug Abuse Messages. Two of the messages that were found to have quite opposite and interesting effects on the parents audience were run in a month long split cable field experiment in the Ad Tel West split cable market. The two messages were run on alternate cables with a 400 GRP weight over a 4-week period in May and early June 1973. The primary finding for the present purposes was that the laboratory findings with regard to the two commercials (which were described in the previous paragraph) held up in the field experiment. The strong but irritating commercial did best on all field measures, with the exception of those measures having to do with advertisement liking. Another finding relevant to those doing anti-drug abuse advertising was that it was possible to have a marked effect on the belief atmosphere in a local community with a heavy saturation campaign.

Studies XI and XII: Television Clutter Research. Peter Webb, with support from the Marketing Science Institute, has done two large-scale studies in which the key independent variable was the degree and nature of television clutter, i.e., the amount of nonprogram interruptions (primarily advertising) and the way they are scheduled within a program. In this research he used the "Television Violence-Humor Project" and a new "Television Program Rating Evaluation Project" cover stories. This research involved an observational measure of respondent attention while viewing the commercials. Respondents were told that they should watch the programs as they normally would, and coffee and doughnuts were provided in the room to get them moving about as they normally would when they watch television. One preliminary finding was that on many of the measures, the gross amount of clutter was not as important as the position of each advertisement within the clutter stream and the scheduling of the nonprogram material during the program itself.

#### Current Perspectives

Developing a pretesting technique is an engineering job. It requires movement back and forth over the stages of the research system outlined here. We believe we have a useful pretesting technique that can be used in a wide variety of situations. Our research thus far has indicated that the response to communication is quite complex. But it can be monitored with a technique such as the one proposed and used here. The data from such a pretesting technique should be usable in the development of media model applications. Our data have been promising thus far and offer hope for improved advertising planning in the future.

#### REFERENCES

- Ray, Michael L., A proposal for validating measures and models in highly competitive decision situations. In Boris W. Backer and Helmut Becker (Eds.), Combined proceedings. Chicago: American Marketing Association, 1972, 475-478.
- Ray, Michael L., in collaboration with Sawyer, A.G., Rothschild, M.L. Heller, Roger M., Strong, E.C., & Reed, J.B. Marketing communication and the hierarchy-of-effects. In Peter Clarke (Ed.), New models for mass communications research. Beverly Hills: Sage Publications, 1973. Pp. 147-176.
- Ward, Scott & Ray, Michael L. Cognitive responses to mass communication: Results from laboratory studies and a field experiment. Paper presented at the meeting of the Association for Education in Journalism, San Diego, August, 1974. (Stanford GSB Research Paper No. 232, September, 1974.)





## RECENT EXPERIENCES IN COPY RESEARCH

Leland E. Ott<sup>1</sup>  
 Grey Advertising, Incorporated

One of the critical issues in copy testing is the attempt to develop a single technique that can measure all aspects of commercial effectiveness (recall, persuasion, communication, and diagnosis). Results from three testing techniques indicated inherent difficulties in designing a single test as recall ideally requires a natural type of environment while the other ones need a forced exposure type of environment for sensitive measurement. A recent innovation in rough film production called Photomatic, consisting of finished audio and still frame photographs, offers promise for making it easier to test alternative executional approaches on critical measures early in the copy development process to select the most promising ones for finished production. Evidence indicates recall results are similar for Photomatic finished executions.

## Introduction: Philosophy of Copy Research

In her 1972 article, Shirley Young pointed out the need to develop improved copy testing methodology tailored to the nature of the copy testing problem and degree of risk involved in the business decision. It was also indicated that the role of copy research should be limited to determining how effective an execution carries out a predetermined strategy and never should be used to evaluate alternative strategies. Her article also defined the following components of a selling message that should be considered in designing a copy test:

- Attention: Flagging enough of the appropriate target customers.
- Communication: Transmitting a clear message about the assets of the product.
- Persuasion: Overall-persuading the prospect that this brand is generally more desirable than other alternatives, and Specific-that this brand is better than others on the strategically important benefits.
- No Negative Diagnosis: Not antagonizing the prospect such that he may be turned off by the message after repeat exposure.

If time or money were not issues, one would always wish to use a testing technique that evaluated all the above aspects of commercial effectiveness. But, since time and money are vital operational constraints, a partial evaluation of one or more measures of effectiveness is often sufficient. For example, one would need to use a total evaluation for a major change in executional approach but would need only a partial evaluation for a pool out or extension of an existing campaign. Through the judicious use of copy research approaches tailored to needs, it is possible to get a much better value for dollar expenditure than by using a standard copy testing procedure that often does not really apply in many situations.

## Total Evaluation Testing Experience

During the past fourteen years, Grey has striven to develop a copy testing technique that could measure all of the selling components within a realistic viewing environment while maintaining high standards of research control. It was our belief that a testing technique measuring all aspects of the commercial could enable us to cross relate the various measures while at the same time saving money over conducting several tests in sequential stages. Up to three years ago, two basic approaches were used. They were:

### In-Home TV Matic

- . Consisted of a test group exposed in-home by rear view projector to a test commercial in a 20 minute film.
- . Clutter limited to one distractor commercial.
- . Control group not exposed to either commercial or program.
- . Day after interviewing on recall, communication, and persuasion measures.

### On-Air Decision

- . Consisted of on-air exposure under natural viewing environment.
- . Cross over experimental design used to balance out effects of program and cities.
- . Control group exposed to program only without commercial.
- . Day after interviewing on recall, communication and persuasion measures.

After reviewing both approaches, it was concluded that neither provided the desired degree of sensitivity or discrimination (Table 1). At the 80 percent confidence level, the TV Matic discriminated in about one half the cases while Decision did less than one-fourth of the time, or little better than chance. There were also the following additional research problems:

- Both techniques tested against an unexposed control so that the effect of the media was being measured as well as the creative content of the commercial.
- The base for recallers was inadequate to provide good communication measures or adequate diagnostic information about the executional elements.

TABLE 1

The Number of Shifts In Overall Persuasion  
Between Test and Unexposed Control Groups  
(80 Percent Confidence Level)

	<u>Vs. Unexposed Control</u>	
	<u>Shifts</u>	<u>No Shifts</u>
In-Home TV Matic (105 Tests)	55%	45%
On-Air Decision (23 Tests)	26%	74%

We also felt the desirability of developing a multiple exposure testing technique since:

- . The work of Grass indicated that it took at least two or three exposures to a commercial before maximum learning and attention occurred.
- . Changes in the TV environment, eg. greater use of 30 second commercials and increase in number of messages and non program elements, indicated manufacturers are relying on repeated exposure to commercials.

In an attempt to overcome these research problems and provide a test more attuned to the TV environment, Multiview was developed. It consisted of:

- . Multiple exposure (two for :60 commercials and three for :30 commercials) in a twenty minute film by rear view projector in home.
- . Five minutes of distractor commercials with one of them exposed the same number of times as the test commercial.
- . Control group exposed to a non test commercial suitable to the business decision (usually the one currently being run).
- . Respondents called the next day and questioned on recall, persuasion, communications, and diagnostic measures.

Multiview was an improvement over our previous technique in that it provided:

- . Increased sensitivity and discrimination as we could make a clear cut business decision in about 70 percent of the cases (Table 2). This is remarkable when one considers that we are eliminating the effect of media weight by testing against an exposed control.
- . Communications and diagnostic measures were much improved over previous techniques due to the multiple exposure.

TABLE 2

Shifts on Overall Attitudes and Strategic Benefits For  
Multiview Copy Tests (80 Percent Confidence Level)

<u>Type of Shift</u>	
Both Overall and Strategic	13 35 22
Overall Only	
Strategic Only	
Inconsistent Overall and Strategic	4 26
Neither Overall Nor Strategic	

Note.-Based on 23 tests using 6 point Likelihood-to-Buy Scale for overall attitudes and 6 Point Excellent-to-Poor Scale for strategic benefits.

However, there was a serious problem with the recall measure. Since the level of recall was highly inflated (mean score of 73), there was no assurance that the technique would uncover a commercial that would score low in an on-air environment.

Based on the Muliview experience, it has been demonstrated that sensitive discrimination measures can be obtained when there is a sufficiently strong stimulus. This stronger stimulus also improves the communication and diagnostic measures. However, since the very act of obtaining these measures destroys the attention or recall measure, it is desirable to develop two testing approaches, one to measure recall in a natural type of on-air environment and the other to measure persuasion, communication, and diagnosis in a forced exposure environment.

#### Testing Rough versus Finished Commercials

One of the considerations in copy testing is the use of rough commercials in order to save the cost of producing finished commercials. It is an area that is receiving widespread attention as a way of not only saving time and cost but also permitting the development of more executional approaches for copytesting.

Since it is necessary for the rough commercial to be a reasonable facsimile of the finished version, it must be an audio-visual stimulus for TV commercials. For example, print ads or storyboards can never adequately portray a TV commercial.

While we have had experience for several years with both rough film and animatic commercials, we have recently been using a new form of rough production called Photomatic, which consists of a finished audio-track and a visual stimulus consisting of about 10-15 still frame photographs for a 60 second commercial. This executional approach is quick to produce (about 2 to 3 weeks) and the cost is roughly comparable to animatic production or about half that of rough film ranging from \$3,000 - \$6,000 depending upon the complexity of the execution.

Our experience with this technique has been limited to explicit types of executions featuring demonstrations and/or slice-of-life stories in the packaged goods field. For commercials that are highly implicit or require name personalities, special scenes, much movement, etc., we would not recommend Photomatic as these types of executions could not be adequately portrayed. In the final analysis, it is a creative judgment as to whether the commercial can be adequately portrayed by Photomatic or other rough form of production.

Although our experience to date is limited to ten executions for which Photomatic and finished tests are available, the results indicate, that Photomatic executions are comparable to finished ones at least as far as recall is concerned (Table 3). By permitting the testing of more executions and reducing the time to produce and test, this makes it now possible to test in rough and utilize these findings to improve the finished version.

TABLE 3

Comparison of Photomatic and Finished Commercial  
Recall Tests (10 Commercials)

	Related Recall Scores	Main Copy Points
Average Score	<u>27%</u>	<u>22%</u>
Average Difference between Photomatic and Finished Version	2.7%	1.7%
Number Times Finished Higher	4	3
Number Times Photomatic Higher	3	4
Same	3	3
Correlation Coefficient	.96	.93

Conclusion: Implications

Through the use of Photomatic film production, it is possible to weed out the least promising executional approaches at an early stage in the copy development program. Each system should be tailored to the needs of the particular product situation by focusing on those aspects of commercial effectiveness that are the most crucial.

For most established products having moderate to high budget levels, persuasion, communication and diagnostic measures should receive priority in developing a copy testing technique. Since there is ample evidence that positive changes in attitudes do precede positive changes in behavior (eg. see Young, Axelrod and DuBois), it is necessary that a commercial be shown to be persuasive on an overall basis or strategic benefits before being accepted in a high risk situation. Communications and diagnosis of negative elements are also needed at an early stage so that modifications can be made.

Recall testing, however, should receive a high priority under the following situations:

- . For new products since the critical needs are to create a high level of initial brand awareness and consumer interest.
- . Low media budget brands since the message cannot receive a high frequency of repeat exposure.

Under these circumstances, it is critical to consider only those executional approaches capable of attracting and holding viewer attention in today's cluttered TV environment.

FOOTNOTES

1. Leland E. Ott is Vice President and Technical Director, Marketing and Research Department at Grey Advertising, Incorporated.

## REFERENCES

- Axelrod, Joel N. Attitude measures that predict purchase. Journal of Advertising Research, 1968, 8(1), 3-17.
- DuBois, Cornelius. Twelve brands on a seesaw. In Proceedings: 13th Annual Conference. New York: Advertising Research Foundation, 1967.
- DuBois, Cornelius. Brand XL study. Paper presented at the 1959 meeting of the American Association of Public Opinion Research, New York: Foote, Cone and Belding, 1959.
- Grass, Robert C. Satiation effects of TV commercials. Journal of Advertising Research, 1969, 9(3), 3-8.
- Young, S. Copy testing without magic numbers. Journal of Advertising Research 1972, 12(1), 3-12.

## COPYTESTING--COMMUNICATION VS. PERSUASION

1

2

Fran Kahn and Larry Light  
Batten, Barton, Durstine & Osborn, Inc.

A variety of copy testing systems are available for pre-testing commercials. Most of these systems can be described as falling into one of two major categories: systems primarily designed to measure communication and systems primarily designed to measure persuasion. Over the years, there has been considerable controversy regarding which methodology offers the best measure of commercial effectiveness--i.e., which can validly predict the sales potential of a commercial. An extensive investigation of traditional testing techniques led to the conclusion that, relative to other measures, communication testing is the most reliable, valid, and practical measure of a commercial's effectiveness. However, communication testing is useful only when it is used as the final step in a disciplined copy development system.

## Introduction: Communication vs. Persuasion

The purpose of advertising is to increase sales. And the purpose of the advertising test--specifically, the commercial test--is to produce a yes or no decision on the ability of a commercial to increase sales before that commercial is released for broadscale, on-air use. The question then becomes: What is the best way to test a commercial's effectiveness? What kind of a test will determine whether a commercial has the potential to stimulate purchases?

The theoretical answer is relatively easy to state. First, we need a technique which is reliable--i.e., one which is reproducible so that we can be confident that the results observed are due to the performances of the commercials being tested rather than random fluctuations in the system. We also need a technique which is valid. A valid technique will provide predictions that both correlate with real world observations and relate to future purchase behavior.

How do we assess whether a system is reliable and valid? Reliability (or lack of it) can be demonstrated without great difficulty once sufficient test-retest data have been collected. Validity, however, poses a problem because of the time, cost and complexity of conducting the type of controlled experiment necessary to verify a technique. How, then, are we to know whether or not a system is valid--whether it will accurately separate the good commercials from the bad?

A variety of methods are currently available for testing commercials. Some use the criterion of communication (brand name and/or copy recall); some use persuasion (increased brand interest). While still others use a combination of both, there is usually heavier reliance on one measure or the other, depending upon the philosophy under which the system or the user of the system is operating. All systems claim to be reliable and all systems claim to be valid. Yet, among members of the advertising community there is certainly no agreement



that all systems are reliable and valid...and therefore equally acceptable; nor is there agreement as to which one system is most reliable and valid.

Therefore, after many years of experimentation and discussion, the debate about commercial testing continues--specifically, which is a more accurate measure of commercial effectiveness: communication or persuasion?

The following pages represent the results of BBDO's comprehensive search to answer the "communication vs. persuasion" question. Our own experimentation with various techniques, combined with the experiences of our Clients, provided us with a considerable bank of primary data (over 2,300 separate commercial tests among more than 460,000 respondents). The purpose of this document is to share some of the findings that resulted from our review of the major types of commercial testing systems and to present some of the conclusions based on these findings.

### How Accurate Are Commercial Testing Systems?

#### Reliability

How accurate are commercial testing systems? By accurate, we mean reliable and valid. Taking reliability first, the question, put simply, is: How sure can we be in a commercial's test score? Will a commercial initially judged as good (or poor), when retested on the same system, again be evaluated as good (or poor)? If not--if a scoring system involves unpredictable fluctuations that result from factors other than differences in the commercials being tested--then the results are highly suspect.

We approached the issue of reliability by first examining our experience with our own testing system. In the past, much of BBDO's commercial testing was conducted via its Channel One facility. This was an on-air system. Test commercials were run in the context of an actual, regularly scheduled TV program which the Agency purchased in one or two cities. Within three hours after the program broadcast, random telephone calls were made and program viewers were questioned regarding awareness and recall of the test commercial(s). In the seven years BBDO used Channel One, 633 commercial tests were undertaken--including 106 retest cases where the same commercial was tested twice. This is one of the industry's largest sources of test-retest information and therefore allows us to look with confidence at the reliability data developed.

In order to assess the reliability of our own Channel One technique, all 633 commercials were spread out on a scale from low-scoring to high-scoring. The result was a normal (bell-shaped) curve with an average awareness score of 33.4%. Next, we examined the 106 commercials that had been tested twice and found that the difference between the first and second test scores averaged 6.4%. In other words, a commercial's second test score, on average, was likely to have varied up to 6.4 points, either above or below its first test score. By taking the test-retest variation of  $\pm 6.4$  points, and superimposing it back on the distribution of all Channel One scores, we were able to mark out an area of uncertainty--the unreliability zone. Figure 1 shows the unreliability zone for the Channel One system--an area extending from 27.0% to 39.8%, around the the 33.4% mean.

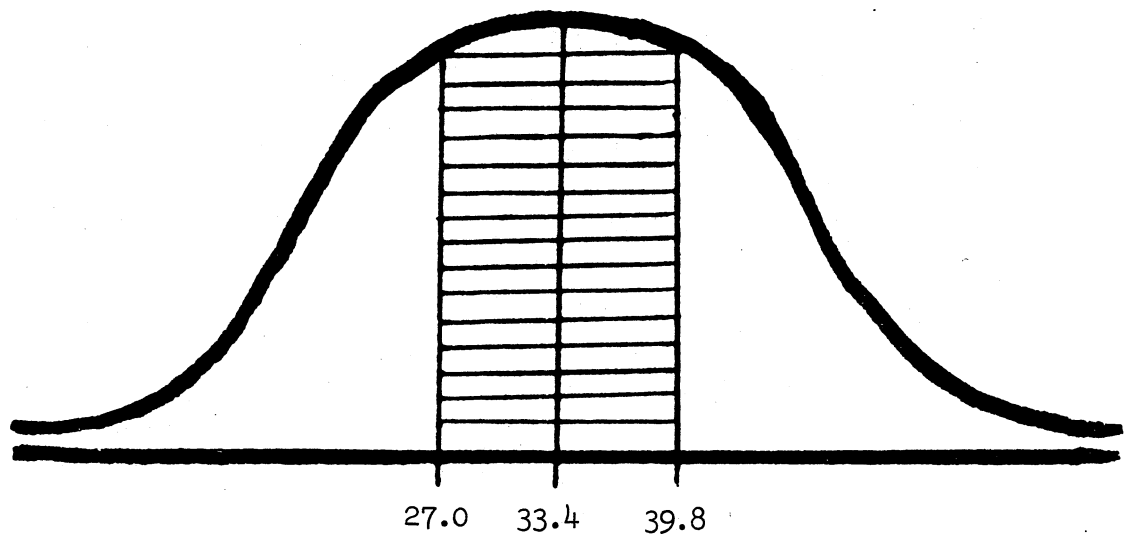


Figure 1. The Channel One Unreliability Zone

What were the implications of this finding? The evaluation of commercials scoring either below or above the 27.0% to 39.8% unreliability zone would not be expected to change when retested. For example, if a commercial scoring 40% on its initial test were retested, and the second score dropped the average 6.4 points, that commercial would still fall on the positive side of the mean. The same reasoning holds for a poor commercial whose original score was lower than 27%. Even with an increase of 6.4 points in a retest, it would still remain on the negative side of the distribution. But a score within the unreliability zone, differing by as much as  $\pm 6.4$  points, could be expected to swing from the "good" side of the mean to the "poor" side in any test-retest situation. Therefore, only those scores falling outside the unreliability zone could be used with confidence for a go/no-go decision. And those commercials--i.e., those falling outside the unreliability zone--represented 50% of all commercials tested. Only 50% of all commercial tests yielded a definitive decision, as is shown in Figure 2.

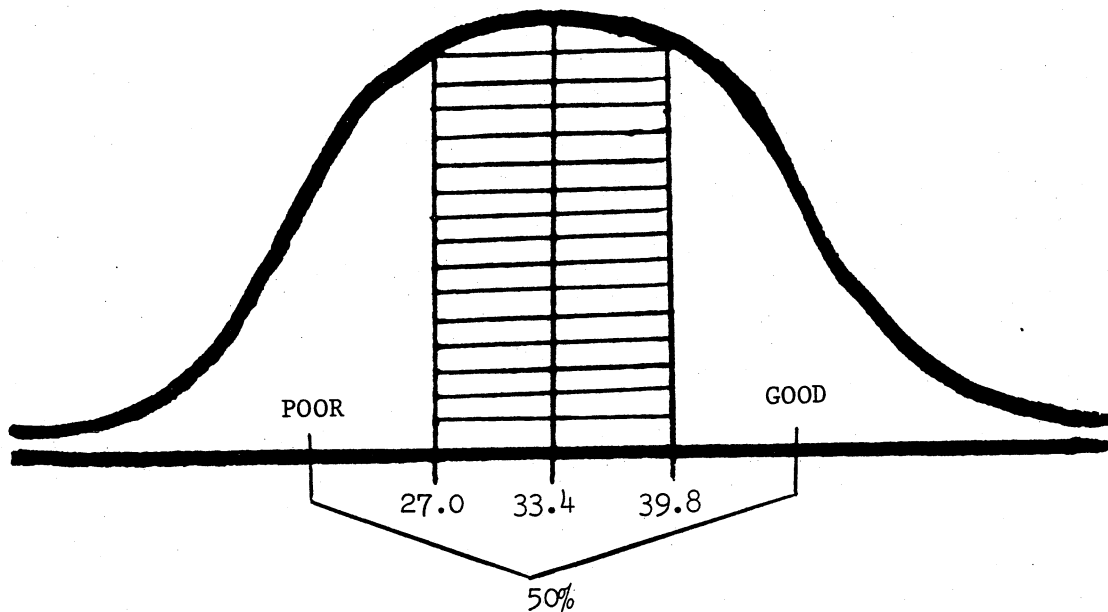


Figure 2. Percentage of Commercials Falling Outside the Channel One Unreliability Zone

How do other testing techniques compare in terms of statistical reliability? Following the reliability analysis of Channel One data, BBDO conducted the same analysis where sufficient data were either supplied by our Clients or available from our own files. We examined two kinds of measures: (1) communication data and (2) persuasion data. The communication data could be further subdivided into conditions of exposure:

- . Natural (on-air), and
- . Forced-exposure (theater)

Similarly, persuasion data could be further subdivided; for while measures of persuasion most commonly utilize forced exposure viewing, one technique we reviewed involved attitude shift (pre to post) and the other used behavior data. The two methods of data collection for persuasion, then, are:

- . Pre-post attitude shift (theater), and
- . Behavior (coupon redemption)

Table 1 summarizes the results obtained after examining various testing techniques for reliability. Also shown in the table is the time interval, used by each system, between exposure of a commercial and measurement of effects.

TABLE 1

## Reliability of Different Testing Systems

<u>Measure/Technique</u>	<u>% of Scores outside Unreliability Zone</u>	<u>Delay Interval between Exposure and Measurement</u>
Communication		
On-air: BBDO's Channel One	50	3 Hours
On-air: System #2	61	24 Hours
Theater: System #3	72	72 Hours
Theater: System #4	46	Immediate
Persuasion		
Theater: System #3	50	Immediate
Coupon redemption: #5	53	Immediate
Coupon redemption: #6	48	Immediate

What did we learn from the reliability analysis? Three very important observations resulted: (1) Commercial systems, in general, have a wide unreliability zone. (2) Some measuring techniques are statistically more reliable than others. (3) The critical difference separating the more reliable systems from the others analyzed is the delay factor between time of exposure to the commercial and time of measurement. Thus, 24-hour on-air recall and 72-hour forced exposure recall can be viewed as the most reliable testing systems of those we studied.

Validity

If we consider that the basic purpose of advertising is to produce sales, then the basic aim of the ideal commercial test is to give the decision-maker an accurate prediction of the selling effectiveness of his commercial. A valid measuring system must meet two criteria:

- . Predictions should correlate with real world observations.
- . What is measured should relate to future purchase behavior.

As a first step, we can get some idea of comparative validity among various systems by investigating the belief that each commercial testing system claims to represent the most valid way of testing commercials. In other words, if all testing systems are valid, then regardless of which is used, a good commercial on one system should also be a good commercial on another and vice versa. Since sales results attributable to commercial exposure are limited, we began by looking at the question of validity in terms of the comparative ability of various testing systems to agree as to how good or bad the same commercial is.

To accomplish this, we collected the experience BBDO and its Clients have had in testing the same commercial on different commercial testing systems. Due to the great variation among the different systems, we classified each commercial's test result in terms of its score relative to the norm for the particular system used to evaluate it. The test data were divided into three levels of classification: commercials characterized as good, as average, and as poor.

Validity: On-air communication vs. theater communication. First we compared results from on-air tests vs. theater tests--using communication as the criterion measure. On-air testing has the advantage of being a real-world observation in that respondents view the commercial in a natural, at-home setting and are unaware at the time of exposure that questioning will occur. The price paid for this advantage is the cost of a finished film (unfinished films may be used in theater testing) plus the media expenses incurred. Our question was: Does the same evaluation occur when two different techniques are used to measure a commercial on a single criterion--communication? Our case histories consisted of 31 observations in which a given commercial was evaluated on two systems--on-air and theater. The result: in 24 out of 31 instances, the evaluation of a commercial on one system was inconsistent with its evaluation on another system. That is, 77% of the time, an advertiser would have come to a different decision about the communication effectiveness of a commercial, depending on the system chosen to aid in the evaluation.

We next analyzed the results of 11 commercials which had been measured for communication within one on-air system (24-hour recall) and one theater system (immediate recall). The correlation of on-air recall to theater recall was +.06. Not very encouraging!

Finally, comparing the results of on-air recall (24-hour) to theater recall when delayed (72-hour), for eight separate commercials, we found a correlation of +.59. Delay, again, seemed an important variable. Its use in the theater test had the effect of bringing the two systems (on-air and theater) closer together.

Validity: Communication vs. persuasion. Since 24-hour communication data from on-air tests are generally inconsistent with immediate recall data from theater tests, we next examined persuasion measures. Theater testing, after all, was developed primarily to provide a reading on the persuasiveness of a commercial. This reading, in most theater testing systems, is the pre-post shift--i.e., the relative ability of a commercial to generate increased interest in a given brand. Thus, proponents of the theater testing system generally acknowledge that memorability of commercial elements under forced exposure conditions may not correlate perfectly with memorability under natural viewing conditions. However, an effective commercial must do two things: it must communicate something about the brand (if only the brand name), and it must persuade. Consequently, communication levels are looked at diagnostically (certain minimum levels of communication are required) and persuasion tends to be regarded as the primary criterion measure.

Realizing the different premise behind each system, however, does not help to answer our basic question. On-air communication techniques claim to be valid and theater persuasion techniques claim to be valid. Our next step, then, was to compare the results of both techniques. Was there any relationship between a high-scoring commercial on one system and a high-scoring commercial on another system? Would both systems agree on what is a good commercial and what is a poor one?

Three separate analyses suggest that the two measures, in fact, produce different conclusions regarding commercial effectiveness.

1. One Client tested each of eight commercials via on-air (to produce a 24-hour recall score) and via theater (to produce an immediate pre-post attitude shift score). The results from the two systems, measured against their respective norms, did show a correlation, but a negative one, of -.60.

2. Next we examined 11 more pairs of tests in which a given commercial was tested for both on-air recall and theater pre-post shift (a different system than the one discussed under Point 1). Here the correlation of on-air recall to pre-post shift was +.26.

3. These findings are very much in line with a study reported by Young (1972) which compared the relationship of 15 recall scores to 15 attitude shift scores. In this case, both sets of data were derived from on-air tests; still, the resulting correlation of +.05 indicated no relationship.

The conclusion: there is no positive correlation between recall and persuasion; the two seem to be measuring different effects.

Validity: Is an on-air persuasion measure feasible? Can a commercial's effectiveness be predicted through the use of an on-air persuasion measure? An on-air measure would meet the criterion of real-world observation and have the advantage of combining two kinds of data: brand preference and communication. Unfortunately, an on-air persuasion measure is truly "easier said than done."

BBDO conducted a study to determine whether a change in reported brand preference could predict the effect of one commercial on future purchase. Overall, approximately 3,500 respondents were questioned about 16 product categories. A test group and a control group were drawn from the same cities. The test group was screened for program and commercial viewership; the control group saw neither program nor commercials. Both groups were probed for (1) brand preference and (2) last purchase. After a three-week interval, respondents were called back and asked about brands actually purchased. A comparison was then made of the brands they originally said they would buy and those they subsequently reported having bought. Only consumers having purchased a brand of a product in the interim between interviews were included in the analytical sample.

The study found the most important factor in determining future purchase to be previous purchase. Eighty-three percent (83%) of future purchases were predicted by previous purchase. Fourteen percent (14%) of the purchases could not have been forecast by either past purchase or brand preference. And, importantly, only 3% of future purchases could have been forecast by brand preference alone. In other words, among the latter 17% who switched brands, the preference measure correctly predicted less than one out of five purchases.

A BBDO Client tested over 100 commercials in an effort to validate on-air preference measures. Results showed that one exposure to an average commercial converted 4% of non-users. The preference measures (both lottery and constant sum were used) again detected less than one out of five of those people who were actually converted to trial of the test brand. What does this mean to the advertiser? If his brand has 20% usage, then one exposure of his commercial will convert 4% of the remaining 80% of households (4% X 80%)--or 3.2% of the households. Using an on-air persuasion test, he will validly predict the change in behavior of one out of five of the converted households (20% X 3.2%)--or 0.6% of the households. To do this reliably would require a sample size of 25,000.

Is an on-air persuasion measure feasible? After reviewing our own experience with traditional measurements, the answer was negative. Consumers bring to any viewing situation their own brand loyalties, brand images, biases, etc. Even if a commercial is remembered, one real world exposure produces very small delayed changes in attitude and/or purchase behavior. These changes, after only one exposure, are too small to measure economically.

Validity and sales. What have we learned so far? Traditional on-air persuasion measures are not feasible--at least not economically feasible. Results from communication tests vary depending on whether the methodology is on-air exposure or forced exposure. Communication and persuasion measures--regardless of exposure conditions--do not correlate. In short, different systems produce different conclusions. And, if two systems come to different conclusions, they cannot both be correctly predicting the same thing--in this case, potential sales effectiveness.

The sales test, under controlled conditions, must therefore be the final arbiter of the "communication vs. persuasion" question. Unfortunately, because of the difficulty, time, and expense of conducting such tests under carefully controlled conditions, there is no massive body of data available to provide a ready answer. However, some case histories do offer noteworthy insights.

One Client attempted to determine the validity of on-air testing. Four commercials were tested for 24-hour recall; two were classified as high-scoring commercials and two were average-scoring efforts. Each commercial was subsequently run, over time, in a controlled experiment and changes in sales were measured. Table 2 shows the results: the high-scoring commercials produced  $2\frac{1}{2}$  to 3 times greater increases in sales than the average-scoring commercials.

TABLE 2

## Relationship of On-Air Recall to Sales

<u>Commercial</u>	<u>On-Air Recall Score</u>	<u>Sales Increase</u>
A	High	+26
B	High	+24
C	Average	+10
D	Average	+ 8

The on-air tests of these commercials also included a persuasion measure, using a pre-post design. As indicated in Table 3, the persuasion measure proved less predictive of sales than the communication measure.

TABLE 3

## Relationship of On-Air Recall and On-Air Persuasion to Sales

<u>Commercial</u>	<u>On-Air Recall Score</u>	<u>On-Air Persuasion Score</u>	<u>Sales Increase</u>
A	High	Average	+26
B	High	High	+24
C	Average	High	+10
D	Average	Average	+ 8

Another case history involves two food commercials--commercials Y and Z. Both were on-air tested for recall, and each was judged to be average. Subsequently, the two commercials were tested via three different theater systems and evaluated against the appropriate norm. On Theater system #1, though both commercials continued to perform equally, they were regarded as high-scoring efforts. Theater system #2 broke the tie; commercial Y, with a low score, was judged inferior to commercial Z which scored high. On Theater system #3, the results were reversed. Commercial Y, with a high score, was evaluated as superior to Z which scored at the average. Not knowing which was the "true" reading, the Client went to a sales test--using commercial Y in one market and commercial Z in another market. The result: no sales increase for the brand in commercial Y's test market and no sales increase for the brand in commercial Z's test

market. As Table 4 shows, on-air recall testing had been predictive of sales in both cases.

TABLE 4

## Accuracy of Testing Systems in Predicting Sales

<u>Testing System</u>	Results for		<u>Sales Predictions</u>
	<u>Commercial Y</u> Average	<u>Commercial Z</u> Average	
On-air			No effect on sales with either Y or Z
Theater 1	High	High	Sales increase for both Y and Z
Theater 2	Low	High	Sales increase for Z only
Theater 3	High	Average	Sales increase for Y only

Outcome of Sales Test: No sales increase for either Y or Z

As final evidence, a study by Bogart, Tolley and Orenstein (1970), examined the relationship between measures of sales response and measures of communication response. Although their experiment used newspaper advertising, the results are relevant to this discussion. Twenty-four (24) ads for packaged goods were measured for sales and recall within a 30-hour period. The conclusion, based on over 25,000 observations, was that respondents who could prove recall reported buying far more of the advertised brand than those who did not read anything. Further, purchase was directly correlated with degree of recall, as shown in Table 5.

TABLE 5

## Purchase of the Advertised Brand, among Those Who Recall It to Varying Degrees

<u>Among Those Who:</u>	<u>Number/1000 Subscribers Who Bought Advertised Brand</u>
Proved recall	9
Recognized ad (when shown)	4
Read something on page other than ad	3
Opened page, read nothing	2
Did not open page	2

Note.-Reprinted from an article by Leo Bogart, B. Stuart Tolley, and Frank Orenstein published in the August 1970 Journal of Advertising Research  
\*p < .002

Taken together, the above data provide a strong argument favoring the validity of on-air communication testing. In all three cases, recall and sales were highly correlated.

What have we learned about copy testing systems? We first saw that delayed communication measures provided the most statistically reliable results. On-air testing is a delayed measurement system. In addition, it is conducted under natural (i.e., real) viewing conditions and it appears to have sales validity.



On-air communication, therefore, seems to offer the advertiser the best predictive tool when it is used as the final step in evaluating commercials before broadscale on-air use.

#### Conclusion: Persuasion and Communication

Are we then concluding that communication is a better tool for sales prediction than persuasion? Yes and no! First, we would suggest that the traditional question of persuasion vs. communication is inappropriate. Experience indicates that the issue requires two questions: (1) persuasion, when? and (2) communication, when? A persuasive idea which is not executed in a memorable and understandable way is as wasteful of advertising dollars as a memorable and understandable execution that says nothing motivating. We need both: persuasion and communication. But, instead of looking for a single, multi-purpose copy testing technique, what is needed is a reliable and valid copy development system. We need not one measure, but a series of systematically-related measures.

A disciplined copy development system is based on a sound philosophy of how advertising works. BBDO's philosophy is that effective advertising memorably communicates to prime prospects that the advertised product or service solves a problem they have with the category. At BBDO, this philosophy has been translated into four copy development steps:

- |  |   |           |
|--|---|-----------|
| 1. Know your prime prospect            |   |           |
| 2. Know your prime prospect's problems | → | Strategy  |
| 3. Know your product                   | → |           |
| 4. Break the boredom barrier           | → | Execution |

The key to this system is knowing that a strategy is motivating before developing advertising executions. Valid procedures exist for accomplishing this objective. Persuasion, i.e., the ability of an idea to generate purchase interest, should indeed be measured. It should be measured each and every time a new concept is considered. But, persuasion should be determined before, not after, copy is written. Persuasion, then, should be measured at the strategy development stage.

Now, the question remains: Has the strategy been accurately translated in the advertising execution? Is the message clear and easy to understand? Several screening techniques are available to tell us which of various executional approaches does the best job in delivering the intended message.

Finally, on-air communication testing is the last step in a total copy development system. Why do we need this last step, if we already know the strategy is motivating and the execution communicates that strategy? We need on-air testing because it provides information on how a commercial performs in the environment in which it is to be used. On-air testing tells us whether a commercial memorably communicates the intended message in the name of the brand.

The validity of any particular technique must be evaluated in the context of the development system within which it will be used. The system that is adopted is a direct function of a company's philosophy regarding how advertising works. While there is considerable variation in corporate philosophies, there is no disagreement regarding what advertising is supposed to accomplish. Effective advertising, regardless of what it says or how it affects attitudes, must ultimately produce increases in sales. How can we discuss validity without discussing sales effectiveness? How can we discuss validity and sales effectiveness without

measuring real world effects? Certain companies are investing in this kind of research--comparing measurement results to real world results. Without this kind of information, BBDO's review would not have been possible.

#### FOOTNOTES

1. Fran Kahn is Associate Research Director at Batten, Barton, Durstine & Osborn, Inc.
2. Larry Light is Director of Research at Batten, Barton, Durstine & Osborn, Inc.

#### REFERENCES

- Bogart, L., Tolley, B. S., & Orenstein, F. What one little ad can do. Journal of Advertising Research, 1970, 10(4), 3-13.
- Young, S. Copy testing without magic numbers. Journal of Advertising Research, 1972, 12(1), 3-12.



THEORY AS A BRIDGE  
 BETWEEN DESCRIPTION AND EVALUATION  
 OF PERSUASION<sup>1</sup>

Clark Leavitt<sup>2</sup>  
 The Ohio State University

This paper reviews an advertising testing procedure that presents the advertising in a simple, uncomplicated way with no attempt to mislead the respondent as to the purpose. However, the response format is extensive--being a set of rating scales reported previously. The scales represent a unique set of dimensions along which consumers map commercials and other ads. The paper deals with the problem of establishing decision rules for acceptance of advertising when independent dimensions must be used as the basis. Data are presented to illustrate recommended procedures.

Earlier publications<sup>3</sup> have described a set of descriptive rating scales for television commercials. Because of their sensitivity and reliability, these scales provided an excellent diagnostic device for comparing both alternative executions and alternative campaigns.

A problem arises, however, in using the scales to provide a more definite evaluation of alternatives. How can the scales be used for quantified predictions of effectiveness? Simply taking the unweighted average of all the scales would give equal weight to each. This arbitrary procedure is hard to defend because certain scales seem to be tapping more important variables than others. For example, most advertisers feel that it is more important to have personally relevant commercials than to have humorous commercials.

Another approach may be to relate the scales to a theory or model of communication. Assuming the root of the problem is the extreme empirical manner in which the scales were originally developed, a theoretical reference point--although it cannot provide a universal set of weights--can at least provide a sense of direction.

A theory that is useful for this purpose is the process model of persuasion used either implicitly or explicitly by most workers in the field. This model asserts that there are four important aspects to the response to a message that mediates behavioral change. These are:

- attention or arousal
- communication of information
- change in attitude or intention
- retention of effects

When these four-part processes are at optimal levels a message is persuasive according to the process model.

In order to relate the eight empirically derived scales presented previously to this theoretical framework, further statistical analyses of the

existing scales were carried out and modifications were made based on these analyses.

#### Method

First, a factor analysis was performed with the objective of achieving the fewest number of factors that would adequately explain the data (that is, simple structure). Four factors were found that produced the best combination of high loadings and low first-order intercorrelation among items on different factors.

The second step was construction of a balanced set of items for two primary factors. Balance was achieved by adding negative loading words to the original set and equating the total number of items on the two factors. This was intended both to strengthen the factor structure and to avoid yea-saying response bias. The best negative/positive balance that could be achieved was 12 positive and eight negative words because the supply of reliable negative words was limited.

#### Results

The first factor to emerge in the new analysis was Stimulating. This includes items from three of the original eight scales: Humor, Vigor and Uniqueness (or Amusing, Energetic and Novel).

The second factor was Relevant. It consisted of three more of the original scales: Personal Relevance, Authoritative and Irritation or Disliked.

The third and fourth factors were the Sensuous and Familiar scales, respectively.

The new Stimulating and Relevant Scales were now modified by adding and/or deleting words so that each consisted of approximately 12 positive and eight negative, words. In addition, the Sensuous scale was broadened by the inclusion of more positive words derived from further testing and a new understanding of this variable. Familiar was left unchanged.

Other changes were made as the result of further testing. The Professional Execution scale seemed unstable and was replaced by words referring to source credibility and realism. These changes were made possible by the fact that the simple structure achieved in the final analysis became a more powerful tool for picking out stable high loading items.

#### Discussion

Comparing results to theory produced several insights. First of all the underlying nature of the first two factors became clear. Stimulation seems to parallel attention and the five scales (see Table I) provide scope sufficient to measure all aspects. It would seem likely that this factor would also relate to wearout since dull commercials could be expected to wear out sooner.

The Relevant factor seems to map the essential condition for change in attitude or intention. It is a good presumption that we intend to use relevant objects and that objects are relevant because we intend to use them.

None of the factors seems to bear on the communication aspect. The word "informative" is part of the scales but correlates with words like "helpful" and "important for me." Obviously for the average consumer irrelevant information is not really information at all.

This means that the scales are not sufficient for checking "communication." Apparently if the analyst wishes to determine whether the copy points were clear (even if not important) he will have to use other means. This is confirmed by actual near-zero correlations with such other measures.

The relation of these four perceptual factors to retention is not yet clear. It is to be hoped that testing the same commercial at various points in its exposure history may help clarify this.

Finally, the greater generality of the sensuous scale comes as a surprise. Theory was helpful here also because the lack of fit made the surprise that much clearer. This scale was renamed Gratifying because experience gained in the studies reported here show that it was elevated not only by low-keyed sensuous commercials but also by characters who were idealized in one sense or another and by themes that were folksy, sentimental or confirmed common prejudice or folk wisdom. Taken together, these stimuli suggest a process of cognitive and emotional closure in the viewer.

TABLE I

STIMULATING	RELEVANT	GRATIFYING	FAMILIAR
<u>Amusing</u>	<u>Personal Relevance</u>	a good world	familiar
amusing	convincing	agreeable	saw before
clever	helpful	attractive	well-known
merry	important for me	dreamy	new
playful	meaningful for me	sensitive	
		soothing	
<u>Energetic</u>	<u>Credible</u>	tender	
exciting	dependable	warm	
fast moving	frank		
lively	wise		
vigorous	worth remembering		
<u>Novel</u>	<u>Realistic</u>		
creative	believable		
imaginative	genuine		
novel	natural		
unique	realistic		
<u>Slow</u>	<u>Irritating</u>		
flat	in poor taste		
dull	phony		
slow	silly		
sluggish	stupid		
<u>Worn Out</u>	<u>Confusing</u>		
copy-cat	confusing		
old	unclear		
repetitious	clear		
worn out	informative		

### Validity

To this point the development of the Consumer Relevance Profile from the original set of perceptual scales has been a classic case of construct validation. That is, the attempt has been to establish a set of postulated variables as self-consistent or reliable. This is a kind of validity since it meets the first requirement of any presumed real variable - that it exists.

The next question is whether these variables are correctly named. Specifically, does Stimulating relate to other measures of attention and does the Relevant dimension predict persuasion or, at least, does it correlate with other--independent--measures of persuasion?

Following are some demonstrations that address the validity questions. The results are brief but encouraging.

These results don't provide a statistical weighting, of course. Another approach to a more exact definition of effectiveness might be by analyzing patterns of scores that seem to indicate poor performance.

For example, in a sample of 5 radio commercials dealing with shoplifting, the one judged least relevant was most stimulating. This is typical of a case where the entertaining aspects of a commercial seem to overpower the persuasive effects.

The opposite case would be a low stimulating score and a high relevance: commonly produced by a simple announcement format. This is not likely to wear well.

A familiar spokesman can produce average relevance but an increase in Gratifying and in Familiar. This could be a danger signal--watch out for an increase in irritation

### Conclusion

Although research here doesn't lead to universally applicable single score, it is moving in that direction. The most general conclusion is that a commercial should be high in both Relevance and Stimulation but not excessively high in one without the other. It may be possible to refine this and possibly quantify it.

### Tracking Wearout

A rather dull commercial for a new cereal product was tested when it was first aired and again a year later. As would be expected in the case of a commercial with a heavy schedule, the three largest differences of all the 12 scales were:

	<u>1970</u>	<u>1971</u>
Irritation	169	225
Worn Out	173	212
Familiar	246	283

Perhaps if the commercial had been more stimulating to begin with it would have increased in Familiar alone and not in Irritation.

A second study compared two commercials for the same product early in

their career on the air and six months later. The retest was done with a theatre audience who checked some of the words used in the Perceptual Scales. Here are the results for words used that were about the same on the original test (or, in a few cases, opposite in direction of response).

<u>Words</u>	<u>First Test</u>		<u>Second Test</u>	
	<u>Falls</u>	<u>Burn</u>	<u>Falls</u>	<u>Burn</u>
Worn out	160	163	15%	24%
Familiar	323	297	34	50
Phony	143	147	22	31

The second test cannot be compared directly to first. However, for these three words the difference in the first test are trivial in two cases and the third case they favor the Falls commercial. The percentage differences are all much larger for Burn in the second test.

This seems to reflect the effects of wearout because Burn was exposed more than Falls and was also quite similar to another commercial for the same product. As in the previous case, here, too, Burn was less stimulating originally and could, therefore, be expected to wear out faster.

#### A Creative Award Commercial

An award-winning commercial was judged to do an excellent job of creating awareness of a product change but seemed to place too little stress on product benefit. A new commercial was produced which dealt almost exclusively with the end benefit of using the product.

	<u>Novel</u>	<u>Personal Relevance</u>
Award winner	292	232
Benefit Oriented	232	328

The new campaign was less novel and creative (less stimulating) but had greater personal relevance. It might be argued that the two commercials performed in tandem with a high degree of effectiveness for introducing a product change.

#### Adapting TV Commercials to Radio

Three different types of television commercials were adapted to radio. Eight radio commercials distributed among the three types were tested to see which kind of adaptation succeeded best. To determine this, listeners were given both the Perceptual Scales and a buying interest question. The three types of execution were different

<u>Type</u>	<u>Personal Relevance</u>	<u>Buying Interest</u>
Dialogue	323	78%
Announcer	301	68%
Drama	270	64%

Commercials using a dialogue were better than the standard announcer technique but those using a dramatic format did less well in producing buying interest and relevance. The complications of the dramatic form apparently are harder to get across on TV than on radio.



### New Product Commercials

Three client commercials and two for a new brand of frozen dessert that was competitive to one of the client commercials were tested by a syndicated service. (McCollum-Spielman)

The service administered a shortened version of the Relevance word list in the form of a check list (respondents checked "yes" or "no"). In addition the audience was asked to fill out the service's standard intention-to-buy scale that was always used with new products.

	<u>Buying Interest</u>	<u>Relevance</u>
Client type A	70	120
Client type B	58	114
Competition type A (new)	58	111
Competition type B (old)	46	97
Client type C	46	90

Results show perfect agreement even though fewer Relevance words were used by the service and the manner of responding was a simple check instead of a five-point rating. New product commercials is an area where the use of an intention-to-buy scale is fairly common and more defensible than with established brands. The fact that buying intention and the Relevance words were in agreement in this situation implies that Relevance might be useful in evaluating new products.

### Corporate Versus Line

Two gasoline commercials were tested: one corporate--and one product-oriented

	<u>Personal Relevance</u>
corporate	297
product	348

The product-oriented commercial has a much higher degree of personal relevance for ordinary consumers. The results might have been different if the test had been carried out on stockholders or employees.

### Theatre Test Compared

The lowest and highest ASI pre-post change scores for 60-second carryout food product commercials were compared. The Personal Relevance averages done with 30 people in the Burnett Laboratory were also lowest and highest for eight 60-second commercials tested.

	<u>Percent Change</u>	<u>Personal Relevance</u>
pre-post high	13	341
pre-post low	1	241

### FOOTNOTES

1. This research was supported by the College of Administrative Science of The Ohio State University.

2. Clark Leavitt is Professor of Marketing at The Ohio State University.
3. Wells, W. D., Leavitt, C., & McConville, M. A reaction profile for TV commercials. Journal of Advertising Research, 1971, 2 (6) 11-17. Wells et. al. report the earliest analysis and Leavitt (1970) a further step in developing the scales. Note the publication dates are in reverse order because of different lags.

#### REFERENCES

- Heyner, K. Expression in music: A discussion of experimental studies and theories. Psychological Review, 1935, 47, 186-204.
- Leavitt, Clark. A multidimensional set of rating scales for television commercials. Journal of Applied Psychology, 1970, 54, 427-429.
- Osgood, C., Suci, G., & Tannenbaum, P. The measurement of meaning. Urbana: University of Illinois Press, 1957.
- Wells, W. D. EQ, son of EQ, and reaction profile. Journal of Marketing, 1964, 28, 45-52.
- Wells, W. D., Leavitt, C., & McConville, M. A reaction profile for TV commercials. Journal of Advertising Research, 1971, 2 (6), 11-17.



1

PREFERENCE TYPES AMONG VIEWERS OF TOOTHPASTE COMMERCIALS

Dwight A. Williams<sup>2</sup>  
University of Missouri at Saint Louis

This study takes an in-depth look at preference types among viewers of televised toothpaste commercials. Four construct elements of style and content (Reality, Moral Conflict, Complexity, and Seriousness) were built into a balanced-block design and used to develop a thirty-six item Q-sort instrument. Each item described a hypothetical TV commercial for a hypothetical brand of toothpaste. Thirty-five people, representing several variables, sorted the items along a modified normal-curve distribution, according to personal preference. A 35 x 35 person-to-person correlation matrix was factor-analyzed. Two strong factors (or preference types) resulted, including: I, "The Entertainment Seeker"; and II, "The Information Seeker."

A number of research techniques have been developed which are suitable for studying individual consumer decision-making on an in-depth basis. Generally, however, those techniques which tell us the most about individual decision-making are limited in their ability to provide us with comparative data about whole classes of consumers. Similarly, the techniques which tell us the most about consumers as groups are limited in the depth of information they can provide. Combining several of these techniques, however, can go a long way toward reaching the ideal of providing in-depth information about whole classes of consumers. This study illustrates one such combination of in-depth research techniques.

The findings reported here were part of a larger study which undertook to compare people's preferences across several media, including television programs, radio programs, movies, magazine articles, and political messages as well as consumer brand commercials (Williams, 1971a, 1971b). Regarding commercials, the study sought to answer specifically (1) what constructs do viewers use in deciding their preferences among commercial messages, and (2) what preference types can be found among viewers of commercial messages?

#### Design

The design used in this study of viewer preferences among commercials is an adaptation of a design which has been developed in recent years to study viewer preferences among television programs. The basic design and its rationale were originally developed by Robert Monaghan (1968a). Several studies by Monaghan and his colleagues illustrate the approach (Harries, 1966; Monaghan, Plummer, Rarick, & Williams, 1974; Monaghan & Titchener, 1969; Plummer, 1968; Rarick, 1967). The objective of these studies is to help the creator of media messages see the world through the viewer's own eyes. Typically, the studies combine such techniques as the focused interview, the repertory grid, and Q-methodology. Each of these techniques by itself provides useful, in-depth information about individual decision-making. Combining these techniques, however, makes it possible to move progressively from the most individualized in-depth information to more systematic comparative information on viewer decision-making. At the end of the chain it is possible to draw some in-depth inferences about the decision-making of those people who are clustered together

into types of viewers.

The instrument used in this study was built from a balanced-block design incorporating four facet elements of style and content, or "appeal elements," found in television programs. This particular set of facet elements was chosen based on earlier studies of viewer preferences. (Monaghan, 1964, 1968b; Monaghan, Plummer, Rarick, & Williams, 1966). These studies used focused interviews (Merton, Fiske & Kendall, 1956), and a form of George Kelly's (1955) repertory grid to elicit constructs and discover the elements of appeal believed to be working most strongly in viewers' decision-making regarding preferences for television programs.

In essence, the facet elements built into the instrument have the characteristics of Kelly's "constructs". Kelly's "personal construct theory" provides a good explanation for individual decision-making processes. Each person has his own "personal construct system" which is made up of superordinate and subordinate bi-polar constructs. The person then applies the appropriate constructs to a particular area of decision-making in his life. Thus, for example, a viewer might apply the construct "Humorous--Serious" in deciding his preferences among television programs or commercials. The viewer then makes a prediction regarding the "humorousness" or "seriousness" of the program, and selects the one which he predicts will be most consistent with his preferred end of the construct. This constant process of making predictions, in turn, helps him refine his construct system. Kelly's repertory grid test provides a way to discover the individual's constructs, and plotting procedures make it possible to graph or "map" his construct system. By providing certain constructs in the grid test, such as "favorite TV program" or "ideal commercial" it is possible to discover which constructs are most important in determining his preferences in any one area.

The important point here is that the process of instrument-building began with techniques which are especially useful for studying individual decision-making in depth. From this earlier data were culled what appeared to be the most important decision-making elements. These in turn were converted into facet elements strongly resembling "constructs", and built into the balanced-block design. The Q-instrument thus contains a hypothesized set of constructs which assisted in the process of analyzing the factor data arrays. Following are the facet elements and their definitions as applied to toothpaste commercials.

#### A. Reality

A1 Believable Facts. The presentation of events as they actually are, or were; persons being presented as themselves. Such phrases as "documentary" or "in his own words" may help distinguish commercials in this category from commercials in the next category.

A2 Believable Fiction. Actors play roles or characters other than themselves, but portray "possible" circumstances. Here, the commercials are staged, actors are used, and conversations are scripted, rather than spontaneous.

A3 Unbelievable Fiction, or Fantasy. Represents improbable, fantastic, or cartoon characters in "believable" circumstances; or "believable" characters in unlikely situations; or a combination of both "unreal" conditions.

#### B. Moral Conflict

B1 Intellectualized Morality. The issues in these commercials cannot be so clearly defined in terms of the struggle between good and bad. The issues

are complex; all is not black or white; there are "shades of gray". Shield helps maintain good dental health, but is not the all-conquering hero.

B2 Sentimentalized Morality. Here the struggle between good and bad is much more pronounced. Shield is clearly the champion of healthy teeth, fighting off the enemies. The main characters in these commercials are likely to be homey, sentimental folks--mother, father, children, school-teacher, etc.

B3 No Moral Conflict. Here the struggle between good and evil is not drawn at either the intellectual or sentimental level. This is simply a message, entertaining or serious, in which the brand is highlighted. This might be a statistical presentation, or a "soft sell" type of commercial.

### C. Complexity

C1 High Complexity. The presentation of material in these commercials is complex and the course of the commercial cannot be so easily predicted. It requires more alertness and attention to follow. Perhaps a series of persons are being interviewed, or several aspects of the product are being discussed.

C2 Low Complexity. The presentation of material in these commercials is less complex and the course of the commercial is more easily predicted. It requires less alertness and attention to follow. By contrast to the above, perhaps only one person or family is being interviewed, or only a single aspect of the product is being discussed.

### D. Seriousness

D1 Humorous. In these commercials, the term "humorous" refers broadly to the "light touch", and may range all the way from "warmth" to "satire and parody". The commercial may feature "warm, friendly people" who may be seen laughing, smiling, or generally enjoying themselves. Perhaps the answers in interviews or conversations may be funny or humorous. The situation, plot, or characters may be funny and have a humorous quality. Finally, the commercial may have the quality of satire or parody.

D2 Serious. The treatment in these commercials is serious, business-like, and straight-forward.

The facet elements were built into a balanced-block design, as shown in Table 1. The design was used to generate factorially 36 possible combinations of these elements (3 x 3 x 2 x 2). Each one of the possible combinations was represented by an item in the instrument. Thus, item #1 combines elements A1, B1, C1, D1, and item #36 combines elements A3, B3, C2, D2. Each item was written as the description of a hypothetical television commercial for a hypothetical brand of toothpaste(Shield). The items were typed on 4 x 6 cards. The instrument was checked by inter-judge panel agreement and refined to be sure that the items represented the elements assigned to them. The Toothpaste Commercial Messages Instrument is presented in Table 2. The coding for each item is included in Table 2, but was not included on the original cards.

Thirty-five people were interviewed. These people represent a wide cross-section of demographic variables and life-styles. Included were persons from all four socio-economic status levels as defined by the 1960 census of Columbus, Ohio. There was an almost even split between the sexes, and about a fourth of the persons are Black. There was a good representation of major age ranges.

TABLE 1  
Block Design of Facet Elements of Style and Content  
Used in Toothpaste Commercial Messages Instrument

Constructs	Facet Elements of Style and Content		
A. Reality	A 1 Believable Facts	A 2 Believable Fiction	A 3 Unbelievable Fiction, Fantasy
B. Moral Conflict	B 1 Intellectualized Morality	B 2 Sentimentalized Morality	B 3 No Moral Conflict
C. Complexity	C 1 High Complexity		C 2 Low Complexity
D. Seriousness	D 1 Humorous		D 2 Serious

TABLE 2  
Toothpaste Commercial Messages Instrument

1. "HOUSEPARTY" (A1, B1, C1, D1). In an Art Linkletter "Houseparty" type of interview, several youngsters are asked to tell what they think causes bad teeth and what they recommend to save teeth. They touch on many subjects-- eating sweets, not brushing, etc. Generally speaking, the interviews are funny and the kids really do "say the darndest things". We are reminded that new Shield toothpaste helps.

2. TEETH AND SURVIVAL (A1, B1, C1, D2). In a series of scenes we are reminded how important good teeth are to animal and human survival. We see pirana fish eating meat, see the powerful teeth of certain wild animals, and see references to the saber-tooth tiger. We even see some unique human uses --such as Eskimo women chewing leather to make it soft and workable. We are then reminded that using Shield toothpaste will help to keep our teeth healthy.

3. THE LITTLE DENTIST (A1, B1, C2, D1). In this "Candid Camera" type scene a little boy is dressed up like a dentist and asked to imagine that he is a dentist giving advice in his own words to a patient on how to take care of his teeth. He is quite a "ham", and in his funny way tells the cause of tooth troubles.

Shield is identified as the sponsor of the message.

4. A SOUTH SEAS BEAUTY (A1, B1, C2, D2). This documentary type film shows a beautiful native South Seas Island girl dancing. Her dancing creates a scene of beauty.

Then, she opens her mouth! The lack of good, healthy teeth, etc. destroys the illusion of beauty.

We are told that "unhealthy teeth are the great destroyer of beauty". And we are reminded that Shield when used in a program of teeth care, can help keep our teeth healthy.

5. BOYS' VIEWS OF SHIELD (A1, B2, C1, D1). A number of school boys each are interviewed about their new toothpaste, Shield. In their own words they talk about its taste, the smaller number of cavities since using Shield, etc. Often, their comments are humorous and the boys themselves appear to be relaxed and having fun. But, through it all, Shield comes through as the champion toothpaste for youngsters who care about having good teeth.

6. MOTHER, TEACHER, DENTIST . . . (A1, B2, C1, D2). Several people important in the life of a growing boy are interviewed and in their own words talk about the importance of having healthy teeth and the results of using Shield. Mother, teacher, dentist, etc. are all interviewed and in their own way end up endorsing Shield as a champion of healthy teeth.

7. A MODEL'S TEETH AND HER CAREER (A1, B2, C2, D1). A beautiful, friendly, talkative young fashion model with a good sense of humor is interviewed about her feelings on new Shield toothpaste.

She tells us in her own often humorous way how important bright, sparkling teeth are to her career. She has been very happy with new Shield. She thinks it has made the difference in her career.

8. A MOTHER'S WORD (A1, B2, C2, D2). A mother, photographed with her son, is interviewed about the toothpastes her son has used. She tells us in her own words how important it is to her that her son have healthy teeth, and how pleased she is that Shield has reduced the number of her son's cavities.

9. HOW DO YOU BRUSH YOUR TEETH? (A1, B3, C1, D1). Here we see candid shots of many different people brushing their teeth. It has a humorous note in that it shows many surprisingly different ways people brush their teeth. . . and the sometimes funny expressions on people's faces. Everyone photographed, of course, is using Shield. We are reminded, "No matter how you brush your teeth--use new Shield!"

10. A TALK WITH THE SCIENTISTS (A1, B3, C1, D2). We see interviews with a number of research scientists who worked at the Dental Research Center which developed the new anti-cavity ingredient in Shield. They talk in their own way about the new ingredient from a scientific standpoint.

11. GIVE IT YOUR OWN TEST (A1, B3, C2, D1). This short commercial focuses our attention on a graph showing how Shield reduced cavities for a test group. In the background we hear an interview with one of the youngsters in the test. The interview is humorous. In the same light note, the announcer asks us to try new Shield--give it your own test.

12. WHAT ARE THE CHANCES? (A1, B3, C2, D2). In this commercial we are presented the stark, simple statistics--what percent of the people will lose their teeth before age 35, what percent by age 45, etc. The question is then asked, "What are you doing to protect your teeth?"

Shield is identified as the sponsor of the message.



13. HOME FROM THE DENTIST (A2, B1, C1, D1). In this take-off on many modern toothpaste commercials, the children have just returned from the dentist and have an unusually good record for reduced cavities this trip. Father is scratching his head trying to figure out what accounts for the remarkable new record. He checks in on the children to see what they're eating--no, not much change there. He checks a couple more leads. Still no clue. Then he checks the bathroom where he sees the new Shield they've been using. "Yep! That must be it!" he says with a wink to the camera. The effect is tongue-in-cheek and humorous.

The announcer chuckles and says, "Well, new Shield alone may not make all the difference--but for the family that really wants to cut down on cavities--it can be a big help!"

14. TEACHER'S WORD (A2, B1, C1, D2). A man representing an authoritative sounding, professional looking teacher is talking to a classroom. He is talking about the many things which contribute to tooth decay and periodontal diseases. Then, he makes the point that regular, proper toothbrushing contributes most to good dental health. And he recommends a toothpaste like new Shield.

15. THE SKELETONS SPEAK (A2, B1, C2, D1). In this tongue-in-cheek commercial, an attractive pair of young college boys are seen talking good-naturedly in a biology lab. While Prof. is out, they begin playing around with a couple of the skeletons in the room. One of the boys is an amateur ventriloquist. They begin manipulating the jaws of the skulls as if they are puppets.

"Notice my bright shiny teeth?" one of the skeletons says.

"Yes! And how do you do it?" the other asks.

"It must be the toothpaste I use," the first voice replies good-naturedly.

The boys laugh.

The announcer assures us it's all in good fun, reminds us that there are several steps to healthy teeth, and suggests that Shield might help after all.

16. A SKELETON'S TEETH (A2, B1, C2, D2). A young man representing a teacher is lecturing to his college class. They are discussing why the teeth of a skeleton survive so much longer than anything else. There is some talk about what makes for good, strong, healthy teeth in the first place, and some suggestions for maintaining good, healthy teeth. We are reminded in the end that new Shield has brought the message and can help.

17. LITTLE JILL'S FOOD (A2, B2, C1, D1). In this humorous staged interview, a mother comments on "how difficult it is to get my little Jill to eat what's good for her--if she doesn't like the taste." The camera passes over all kinds of recommended, but not especially tasty foods. Then she adds, "And you know--I used to have trouble getting Jill to brush her teeth--until new Shield came along. Didn't like the taste of her old toothpaste. But Shield changed all that."

18. "THAT'S IT!" (A2, B2, C1, D2). In this "slice of life" drama, Mom and Dad are excitedly happy over their 4 children's latest dental check-up. It seems they have fewer cavities. The talk centers around what could have caused the improvement--and the several benefits in other ways resulting from the improvement. Finally, like a surprise ending, Dad says, "Golly! Do you think it could be the new Shield we switched to?" Mom says, seriously, "Why, of course. That's it!"

19. A STAR IS BORN (A2, B2, C2, D1). The rise to stardom of an aspiring young actress is traced in this tongue-in-cheek commercial. A point is made

of her choice of a toothpaste. Then, the announcer tells us, humorously, that "Shield alone may not make you a star. But when looks are important to your career, new Shield can be a strong friend in the right place--something like having a good agent."

20. A TIMELY TIP (A2, B2, C2, D2). Here is the short story of an aspiring young model who was having a hard time breaking into the big league of modeling. A top photographer seriously gives her the tip to "try new Shield." She does. And then we see her soon afterwards being photographed for the top magazines in the country. She tells us she has new Shield to thank for her success.

21. A HEALTHY, HAPPY WORLD (A2, B3, C1, D1). This message consists of a series of short scenes of happy, healthy people doing things--usually in some way involving their mouth. They are eating, talking, laughing, smiling, playing sports, partying, even kissing. The message ends with the phrase, "Healthy people have more fun in life."

Shield is briefly identified as the sponsor.

22. WHAT'S NEW IN SHIELD (A2, B3, C1, D2). In a series of interviews, actors representing research scientists in the laboratory tell us about the new ingredients in Shield toothpaste. They tell the story of how these new ingredients were discovered and what they are intended to do for your teeth.

23. A VIBRATING ELECTRIC TOOTHBRUSH (A2, B3, C2, D1). In this early-morning toothbrushing scene, a young working man is sleepily trying to get toothpaste onto his new vibrating electric toothbrush. He has all sorts of difficulties--he misses his aim, he squeezes too much, it falls off the vibrating brush, etc.--finally he makes it. A big grin breaks over his face. His toothpaste, of course, is new Shield. The voice-over says, "New Shield--a good way to start the day."

24. THE BIG PICTURE (A2, B3, C2, D2). A man representing a teacher is seen showing a group of grade-school children a large picture of a tooth. He explains the different parts of the tooth and tells why brushing is important.

Shield is briefly identified as the sponsor of the message.

25. PUT UP A SHIELD (A3, B1, C1, D1). In a sequence of cartoon scenes, several different villains of the vaudeville type appear, each representing a different one of the several causes of bad teeth, such as tobacco stain, lack of time to brush after every meal, etc.

The sound track reads something like, "That old villain tobacco stain trying to dull your healthy teeth? Put up a Shield, and defend your good teeth."

Then, a cartoon shield pops up, and the villain is foiled.

26. THE VIEW FROM INSIDE (A3, B1, C1, D2). An actor representing a scientist is telling us about the several causes of tooth decay. Through a bit of film magic, we are able to go inside the mouth and the teeth appear man-sized. Through speed-up techniques, we see the process of decay in action. The scientist then tells us that proper brushing is one of the ways that help.

Shield is briefly identified as the sponsor of the message.

27. A PIRANA'S ADVICE (A3, B1, C2, D1). Two live pirana fish, with their teeth glistening, appear to be having a conversation. The dubbed-in voices provide a conversation about how important teeth are, and ways to keep them healthy. The announcer reminds us that, "Whenever teeth are important--you find folks using new Shield toothpaste."

The treatment is tongue-in-cheek humorous.

28. THE STRENGTH OF GIBRALTER (A3, B1, C2, D2). The scene opens with a picture of Gibraltar. The announcer says, "Imagine that Gibraltar--were a tooth." At this point, Gibraltar appears transformed into a giant tooth. Then, the comparison is made about erosion and decay, etc. Finally, the commercial ends with the theme--"To give your teeth the strength of Gibraltar--use new Shield."

The treatment is serious throughout.

29. THE FANTASTIC VOYAGE (A3, B2, C1, D1). In this take-off on the Fantastic Voyage kind of movie, a dentist and his assistants are inside a cavernous mouth trying to save a tooth. The experience is full of surprises and dangers --but the dental party finally concludes that the only thing that will help now is new Shield. They want to tell this to the patient--if they can ever get out!

30. THE QUEEN OF LIGHT (A3, B2, C1, D2). In a series of scenes, Glisten-anna, the Queen of Light and Brightness, solves several different crises of love by brightening up dull teeth. She tells us seriously we can do the same with new Shield toothpaste. All the people except the Queen appear real rather than as cartoon or fantasy characters.

31. GOLDILOCKS (A3, B2, C2, D1). In a cartoon take-off on "Goldilocks and the Three Bears" we see Goldilocks about to brush her teeth after furnishing the porridge and before she goes to bed. She tries three brands of toothpaste --Pappa Bear's, Mamma Bear's, and Baby Bear's.

"This one tastes too bad," she says of the first.

"And this one stings too much," she says of the second.

"Ah! But this one is just right," she says of the third.

The third, of course, is new Shield.

32. CINDERELLA (A3, B2, C2, D2). Cartoon characters enact a "Cinderella" type story in which Prince Charming falls in love with and asks Cinderella to marry him. He confesses that he was overcome by her smile. At the end, we're reminded how much beautiful teeth can help your smile--and "Who knows--new Shield may even help you get a smile that will conquer a prince."

All in all, the treatment is one of fantasy and romance, charmingly presented. And it is presented as a serious message.

33. SHIELD AND THE SEVEN DWARFS (A3, B3, C1, D1). In a cartoon story with a Walt Disney quality, the Seven Dwarfs are seen marching down the trail from the castle-like Shield toothpaste factory in the background. With toothbrushes slung over their shoulders, they march along singing, "Hi Ho, Hi Ho, A-Brushing we will go-o-o!" In a series of comical scenes, they stop along the way to help brush the teeth of various animal characters in this enchanted forest. Especially funny is when Dumpy tries to help the frog--bubbles everywhere! The treatment is charming and humorous.

34. A TOUR OF THE FLOURIDE PLANT (A3, B3, C1, D2). In this cartoon, a scientist in a lab coat takes us on a tour of the plant where the flouride active compound is made for new Shield. It is an interesting and educational tour.

35. THE PAINTERS (A3, B3, C2, D1). The scene opens on what appears to be a pair of painters in their white cover-all uniforms who appear to be painting a strip down the center of a highway using a painting machine. On closer look, however, we see that the painting machine is laying down a big round strip of toothpaste--the color of new Shield. The characters are exaggerated and are "cutting up" quite a bit. We are then told it would take a continuous strip of toothpaste x number of inches in diameter circling around the world x number of times to equal the total output of Shield toothpaste in a year.

"That's how popular new Shield is--worldwide."

36. LAKE SHIELD (A3, B3, C2, D2). Through the magic of the camera, we see a small lake being filled up with new Shield toothpaste. By the end of the commercial, the lake appears to be filled up. The announcer says, "Every year, Shield makes enough toothpaste to fill a small lake. That's enough to brush an awful lot of people's teeth. Why don't you try new Shield and see why it's the world's largest selling toothpaste."

Each person sorted the 35 toothpaste commercial description cards into nine piles. The nine piles represented a rank ordering along the quasi-normal curve distribution from "Most Preferred" to "Least Preferred", as follows:

Pile number & score	Most Preferred					Least Preferred			
	1	2	3	4	5	6	7	8	9
Frequency	2	3	4	5	8	5	4	3	2

Instructions preceding the sort stressed the importance of expressing personal preference. Following the sort, the person was asked to comment at length on the reasons for most preferring the two items in pile number 1, and least preferring the two items in pile number 9.

The sorting of these commercial description items by each individual formed the basis for clustering persons together into "preference types". Each person was correlated with every other person in the sample, using Pearson  $r$  computations. A 35 x 35 person-to-person correlation matrix was built, and then factor-analyzed using a principal axes-analysis with varimax rotation. Only those factors which accounted for seven percent or more of the total variance between people were reported.

Each factor represents a hypothetical type of viewer. This hypothetical viewer is defined according to the shared preferences for toothpaste commercials. In essence, the factor clusters together those persons who share similar "most preferred" and "least preferred" toothpaste commercials, and whose shared preferences were different from all other types. A factor data array was built to show the "best estimate" of the order of preference for items by each type of viewer. Persons who loaded .70 or higher on the factor were considered to be most representative of the preference type. Item scores for the commercials in the factor data array are averages of the scores (rankings) of the items by those persons loading highest on the factor. Inferences about each type were drawn from the data array for that type. Each type was then described in terms of the toothpaste commercials they most accepted and rejected in expressing their preferences.

The balanced-block design, Q-sorting technique, and factor analysis of matrixes of persons are all features of Q-methodology, which was developed by William Stephenson (1953). Q-methodology offers several important advantages for in-depth study of viewer/consumer decision-making. The Q-sorting technique allows the viewer to express his preference for particular commercials in relationship to all other commercials available to him. This is a process of decision-making which is more realistic than techniques which require him to express degrees of liking or disliking of individual commercials without reference to other commercials. The technique of factor analyzing matrixes of persons makes it possible to discover factors, or types of viewers based on

shared decision-making and preferences rather than purely demographic characteristics. The use of the balanced-block design makes it possible to build constructs into the instrument. This in turn assists the process of interpreting the preference patterns of the types in depth. Finally the process of drawing inferences from the factor data arrays is itself a creative, hypothesis-generating process which can lead to important insights which might be missed using traditional experimental designs. Stephenson calls this process "abduction". Stephenson (1961) says, "Abduction is what one does in guessing or inventing, or proposing a theory or explanation or hypothesis: it is the initial proposition to explain facts . . . the emphasis is on the discovery of hypotheses, not their deduction from postulates". Abduction is a process of inference which is very appropriate to in-depth study of decision-making.

### Results

The sorting of the Toothpaste Commercial Messages Instrument produced two main factors. Each of these factors represents the preference patterns of two types of viewers of toothpaste commercials on television. Factor I has been labeled the "Entertainment Seeker Type", and Factor II the "Information Seeker Type". Each of these viewer preference types is described below in greater detail. Titles and numbers in parentheses refer to descriptions of toothpaste commercial messages found in Table 2.

Factor I: The Entertainment Seeker Type. The data array for Factor I (Table 3) indicates that this type of viewer prefers those toothpaste commercials which have an entertainment value. He most prefers the elements of unbelievable fiction (fantasy) and comedy in commercials. He rejects the elements of believability, high complexity, and seriousness.

The most prevalent element in the Entertainment Seeker's favorite commercials is comedy. Eight of his twelve most preferred commercials are coded comedy (D1). The comedy ranges all the way from candid views of human foibles to spoofs on commercials and contemporary culture. Commercials illustrating this quality include: (9) How do you Brush Your Teeth?; (27) A Pirana's Advice; (29) The Fantastic Voyage; (35) The Painters; (23) A Vibrating Electric Toothbrush; (19) A Star is Born; (31) Goldilocks; and (33) Shield and the Seven Dwarfs.

The Entertainment Seeker also prefers that his commercials contain the element of fantasy. Seven of the twelve favorite commercials are coded unbelievable fiction, or fantasy (A3). In some cases, the fantasy is achieved through film magic, and in others, the fantasy is in the form of cartoons. The fantasy commercials include: (27) A Pirana's Advice; (29) The Fantastic Voyage; (28) The Strength of Gibraltar; (35) The Painters; (31) Goldilocks; (33) Shield and the Seven Dwarfs; and (34) A Tour of the Flouride Plant.

This preference for commercials which combine the elements of fantasy and comedy indicates a desire to be entertained and amused while being persuaded. This preference for entertainment is apparent from looking at the coding and themes of the commercials which the Entertainment Seeker rejects. He rejects commercials which are based in "believable" reality (whether fact or fiction), which are highly complex, and which take themselves and their messages terribly seriously. In addition, many of his rejected commercials feature children and domestic family settings.

Eight of his nine least preferred commercials are coded believable (A1) and (A2). In some cases these are "believable fact" messages which have a documen-

TABLE 3  
Toothpaste Commercial Factor I:  
The Entertainment Seeker

Item	Code	Title	Score
9.	A1, B3, C1, D1	How do you Brush <u>Your</u> Teeth?	1
27.	A3, B1, C2, D1	A Pirana's Advice	1
29.	A3, B2, C1, D1	The Fantastic Voyage	1
28.	A3, B1, C2, D2	The Strength of Gibraltar	1.5
35.	A3, B3, C2, D1	The Painters	1.5
23.	A2, B3, C2, D1	A Vibrating Electric Toothbrush	1.6
2.	A1, B1, C1, D2	Teeth and Survival	2
4.	A1, B1, C2, D2	A South Seas Beauty	2
19.	A2, B2, C2, D1	A Star is Born	2
31.	A3, B2, C2, D1	Goldilocks	2
33.	A3, B3, C1, D1	<u>Shield</u> and the Seven Dwarfs	2
34.	A3, B3, C1, D2	A Tour of the Flouride Plant 2	2
10.	A1, B3, C1, D2	A Talk with the Scientists	8
11.	A1, B3, C2, D1	Give it Your Own Test	8
20.	A2, B2, C2, D2	A Timely Tip	8
25.	A3, B1, C1, D1	Put up a <u>Shield</u>	8
18.	A2, B2, C1, D2	"That's It!"	8.3
17.	A2, B2, C1, D1	Little Jill's Food	8.5
13.	A2, B1, C1, D1	Home from the Dentist	8.6
14.	A2, B1, C1, D2	Teacher's Word	8.6
6.	A1, B2, C1, D2	Mother, Teacher, Dentist	9

tary quality to them, while others are the realistic "believable fiction" mes-

sages. Included among these are (6) Mother, Teacher, Dentist; (14) Teacher's Word; (13) Home from the Dentist; (17) Little Jill's Food; (18) 'That's It!'; (20) A Timely Tip; (11) Give it Your Own Test; and (10) A Talk with the Scientists.

Seven of the nine rejected commercials are coded high complexity (C1). In these commercials that coding has been used to refer to messages which are more difficult to follow and filled with a number of scenes or people. The high complexity commercials include: (6) Mother, Teacher, Dentist; (14) Teacher's Word; (13) Home from the Dentist; (17) Little Jill's Food; (18) 'That's It!' (25) Put up a Shield; and (10) A Talk with the Scientists.

The Entertainment Seeker also rejects serious messages. Five of the nine commercials he least preferred are coded serious (D2). Several of these feature serious minded professional people, such as scientists, dentists, and teachers. Almost always, the characters in these commercials treat their message quite seriously. These commercials include: (6) Mother, Teacher, Dentist; (14) Teacher's Word; (18) 'That's It!'; (20) A Timely Tip; and (10) A Talk with the Scientists.

In summary, Factor I wants to be entertained. He prefers fantasy to reality. He prefers fun to seriousness. He prefers to see a single idea entertainingly presented, rather than a series of ideas treated seriously in one hard-to-follow commercial. It will take something new and different, something fantastic and amusing to appeal to this viewer.

Factor II: The Information Seeker Type. The data array for Factor II (Table 4) indicates that the Information Seeker most prefers commercials which are believable and do not contain a moral conflict, and that he least prefers commercials which contain an element of sentimentalized morality. The Information Seeker wants to be given the facts about a problem and information about a product which will help solve the problem. The presentation should be believable, and the product should not be presented as a knight in shining white armor who is going to miraculously solve the problem.

Four of the six commercials the Information Seeker most prefers are coded believable, (A1) and (A2). Three of these are believable fact presentations, with a live, spontaneous, or documentary quality to them, including: (1) "House-party"; (10) A Talk with the Scientists; and (11) Give it Your Own Test. The fourth is a believable fiction presentation: (24) The Big Picture.

Four of the six most preferred commercials are coded no moral conflict (B3). Generally, this means that the commercial contains no struggle between good and bad, whether of the intellectual "shades of gray" variety, or the sentimental "good guys versus bad guys" variety. Generally, a mere identification of Shield as the sponsor of the message, or a good-natured invitation to try Shield suffices. Commercials in which moral struggle is conspicuously absent are (10) A Talk with the Scientists; (34) A Tour of the Flouride Plant; (11) Give it Your Own Test; and (24) The Big Picture.

Consistent with the Information Seeker's preference for commercials which are believable and which do not contain a moral struggle is his rejection of commercials which contain the element of sentimentalized moral conflict. Five of the eight least preferred commercials are coded sentimentalized morality (B2). These are the commercials which clearly promote Shield as the good guy and champion of healthy teeth. Commercials containing this element include: (32) Cinderella; (20) A Timely Tip; (17) Little Jill's Food; (30) The Queen of Light; (29) The Fantastic Voyage.

TABLE 4

## Toothpaste Commercial Factor II: The Information Seeker

Item	Code	Title	Score
1.	A1, B1, C1, D1	"House Party"	1
10.	A1, B3, C1, D2	A Talk with the Scientists	1.5
34.	A3, B3, C1, D2	A Tour of the Flouride Plant	1.5
11.	A1, B3, C2, D1	Give it Your Own Test	2
24.	A2, B3, C2, D2	The Big Picture	2
31.	A3, B2, C2, D1	Goldilocks	2
4.	A1, B1, C2, D2	A South Seas Beauty	8
23.	A2, B3, C2, D1	A Vibrating Electric Toothbrush	8
29.	A3, B2, C1, D1	The Fantastic Voyage	8
30.	A3, B2, C1, D2	The Queen of Light	8
17.	A2, B2, C1, D1	Little Jill's Food	8.5
20.	A2, B2, C2, D2	A Timely Tip	8.5
13.	A2, B1, C1, D2	Home from the Dentist	9
32.	A3, B2, C2, D2	Cinderella	9

In summary, it is a combination of the No Moral Conflict--Sentimentalized Morality construct with the Believable--Unbelievable construct which distinguishes the Information Seeker. The sentimentalized morality element appears to be rejected in part because it is unbelievable. The information Seeker wants information presented in a believable, non-moralistic fashion.

## Discussion

The combination of techniques used in the design of this study illustrates one possible approach to studying viewer/consumers in depth. The study built upon earlier work using focused interviews and repertory grids. From these earlier studies were culled a set of facet elements which have the characteris-



tics of constructs. These elements were built into a block design and used to construct a Q-instrument containing descriptions of hypothetical commercials. Viewers sorted these descriptions according to personal preference. Persons were then coorelated with each other, and the resulting matrix was factor-analyzed, producing two strong factors, or preference types. The hypothetical set of constructs which were built into the instrument aided the process of "abduction", or drawing inferences, from the factor data arrays. This basic design could be used to study preferences in other categories of commercials, or to study other aspects of consumer decision-making.

The preference patterns revealed by the two strong factors developed in this study suggests possible creative strategies for creators of brand commercials. Factor I, the Entertainment Seeker, clearly prefers commercials with an entertainment quality. He likes comedy and fantasy, while rejecting realistic, complicated, and serious commercials. By contrast, Factor II, the Information Seeker, prefers commercials which provide him with information about the problem and the product in a believable, non-moralistic format.

#### FOOTNOTES

1. The author is grateful to Robert R. Monaghan for advice in the design and execution of this study, and to the Ohio State University for use of its computer facilities.
2. Dwight Williams is Assistant Professor of Speech at the University of Missouri in Saint Louis.

#### REFERENCES

- Harries, T. E. The systematic creation of a new television program. Unpublished master's thesis, Ohio State University, 1966.
- Kelly, G. A. The psychology of personal constructs. 2 vols. New York: W. M. Norton, 1955.
- Merton, R. K., Fiske, M., & Kendall, P. L. The focused interview. New York: The Free Press, 1956.
- Monaghan, R. R., Television preference and viewing behavior. Unpublished doctoral dissertation, Michigan State University, 1964.
- Monaghan, R. R. Creative strategies in audience analysis. Educational Broadcasting Review, 1968, 2, 29-37. (a)
- Monaghan, R. R. A systematic way of being creative. Journal of Communication, 1968, 18, 47-56. (b)
- Monaghan, R. R., Plummer, J. T., Rarick, D. L., & Williams, D. A. Recommended target audience and appeal elements for the Girl from U. N. C. L. E. Unpublished report, Ohio State University, 1966.
- Monaghan, R. R., Plummer, J. T., Rarick, D. L., & Williams, D. A. Predicting viewer preference for new TV program concepts. Journal of Broadcasting, 1974, 18, 131-142.
- Monaghan, R. R., & Titchener, C. EBR readership profile analysis. Educational Broadcasting Review, 1969, 3, 31-42.

- Plummer, J. T. Audience research in television program development. Education-  
al Broadcasting Review, 1968, 2, 23-30.
- Rarick, D. L. Predicting viewer preferences for visual appeals in television programs. Unpublished master's thesis, Ohio State University, 1967.
- Stephenson, W. The study of behavior: Q-technique and its methodology. Chicago: University of Chicago Press, 1953.
- Stephenson, W. Scientific creed--1961. The Philosophical Record, 1961, 11, 1-25.
- Williams, D. A. Mass media preference patterns: A cross-media study. Unpublished doctoral dissertation, Ohio State University, 1971 (a).
- Williams, D. A. Mass media preference patterns: A cross-media study. Paper presented at the Annual Convention of the International Communication Association, Phoenix, Arizona, April, 1971 (b).



A CORRECTIVE ADVERTISING STUDY<sup>1</sup>

Harold H. Kassarian  
University of California, Los Angeles

Cynthia J. Carlson  
Cornell University

and Paula E. Rosin  
Smith College

In recent years the concept of corrective advertising has become an issue of considerable concern to advertiser, academic, and government decision maker. Unfortunately, the number of research studies on the topic has been precious few. This paper attempts to expand previous findings and further test the effects of corrective advertising in an experimental study. The results support previous research -- that exposure to a corrective advertisement does tend to undo some of the attitudes created by misleading, deceptive and unfair advertising.

Corrective advertising generally refers to a Federal Trade Commission order that a firm not only cease and desist from its deceptive advertising, but also refrain from future advertising of a product for a specified period of time, unless it corrects the misleading impression that its prior false advertising may have created (Anon., 1972). The belief is that the issuance of a conventional cease and desist order is ineffectual and does not help dissipate deceptive impressions that may linger in the mind of the consumer.

The concept of corrective advertising was first formally presented in May, 1970 to the FTC by a group of law students organized under the acronym SOUP -- Students Opposing Unfair Practices, Inc. The students attempted to intervene as representatives of the public interest in the Campbell Soup Co. case. The proceedings concerned Campbell's practice of adding marbles to its soups before showing them in television commercials such that the solid ingredients were displaced, giving the product a deceptively rich appearance. The Commission denied SOUP permission to intervene and Campbell agreed to a simple cease and desist order. Four months and one chairman later, a majority of the commission on its own initiative endorsed the concept of corrective advertising (Anon., 1971).

In rapid succession the staff of the Federal Trade Commission asked for corrective advertising in cases of national scope such as Firestone Tires, Hi-C Fruit Drink, Domino Sugar, Wonder Bread, and Standard Oil's Chevron F-310, in addition to other cases of deceptive local and regional advertising.

In August of 1971, the FTC issued its first final order requiring corrective advertising. The order was directed against the ITT Continental Baking Co., marketers of Profile Bread. The product had been advertised as having fewer calories than ordinary bread and as being of weight reducing value. The FTC charged that in fact the bread had fewer calories per slice only because it was more thinly sliced. The uncontested order required that ITT cease and desist all advertisements for Profile Bread unless not less than 25% of expenditures (excluding production costs) for each media in each

market be devoted to advertising stating that Profile Bread is not effective for weight reduction, contrary to possible interpretations of prior advertising (Anon., 1972).

The assumption that when a product is heavily advertised over long periods of time there may be a deep and long lasting impact on the attitudes, beliefs and behavior of consumers is nothing new to the consumer researcher. Topics such as changes in purchase behavior and purchase intentions, demand for specific products and brands, cognitive reorganization of opinions and beliefs, learning and recall of message content and company image, etc., have long been studied. Further, we have good evidence of the lingering effects of a learned behavior pattern, recall of message content, carry-over effects of advertising, and the forgetting or decay of advertised messages.

Lack of decay is not a rare occurrence in advertising. In a splendid paper by Ray, Sawyer and Strong (1971), all of their 36 analyses indicated some degree of non-decay and two-thirds of their learning curves showed a substantial lack of decay. The tendency toward decay was more evident in such routine measures as advertising recall, but in measures of attitude change, beliefs, ratings of the product, brand preference, and usage, the lack of decay was quite pronounced. In fact, for some measures, in some situations, not only can we expect a lack of decay but actually an increase rather than a decrease of a learned response after an advertising campaign has ended and the consumer is no longer exposed to the message. This finding, coined the sleeper effect by Hovland, Lumsdaine and Sheffield (1949), has been extensively researched and superbly reviewed by Capon and Hulbert (1973). These authors point out the rather respectable research tradition supporting this view and conclude that, although there may be some doubt of a generalized sleeper effect, it is detectable for certain subsets of the population.

The Ray interpretation of the sleeper effect is that, "the advertising itself becomes noxious to the audience after a certain point. Up to this point, advertising helps to hold attitudes or may actually help to improve attitudes. After this point, the ad exposures may actually get in the way of favorable opinions of the product or brand" (Ray, Sawyer and Strong, 1971, p. 16). Thus, when an advertising exposure has ended, the long term favorable response continues undiminished by the unfavorable responses caused by excessive and annoying repetition of the advertisement.

Whether or not this interpretation is correct, the available data point out that the effects of some messages, under certain conditions, are extremely difficult to eradicate, and, in fact, the mere stopping of exposure to advertising may not only fail to cause rapid decay, as assumed by cease and desist orders, but may in fact cause an increase in the level of material learned from the advertised message.

Such data as the Ray, Sawyer and Strong findings are, of course, heavily supported by hundreds of studies in experimental psychology on the extinction of a learned response. Most psychologists, in fact, believe that once something is learned it is never truly forgotten. That is to say, a response tendency cannot be extinguished either by non-reinforcement or non-exposure alone. For a response to be extinguished, the animal or the intelligent consumer must be carefully taught to unlearn the response or to replace the learned response with another learned response.

An uncontested commission order for Hawaiian Punch illustrates how such conclusions gleaned from learning theory can be applied to corrective advertising. The complaint alleged that, through featuring fresh fruits and fruit trees prominently in television commercials and through the use of the phrase "seven natural fruit juices", advertisements had represented that Hawaiian Punch beverages consist predominantly of natural fruit juices. In fact, the complaint said, the predominant ingredients are water and sweetening agents which were added to fruit juice and other ingredients. In addition to banning misrepresentation of natural fruit juice content, the order prohibited for a one-year period any advertisement or label which depicted fruit or juice unless the total percentage of single strength fruit juice concentrate contained in a serving was clearly and conspicuously disclosed, or unless the product contained 100% single strength fruit juice.

This provision in the order was to remain in effect until the firm submitted to the Commission a survey on consumer perception of Hawaiian Punch fruit juice content. The form of the survey was included in the order. The affirmative disclosures were not to be required after the one-year period if 67% of current purchasers of fruit flavored beverages, or 80% of current or prospective purchasers of Hawaiian Punch product, or 95% of current purchasers of these products thought that Hawaiian Punch contained no more than 20% natural fruit juice (Kassarjian, 1974).

In this order, the R. J. Reynolds Foods Company was not only asked to cease and desist alleged misrepresentations, but it had to disclose the true facts until such time as a substantial proportion of consumers were no longer misled into thinking that Hawaiian Punch contained major amounts of "seven natural fruit juices", i.e., until a true process of unlearning, decay or extinction and relearning had successfully occurred. The commission apparently felt that consumers must be carefully exposed to the fact that what they learned earlier was in error, lest they continue to suffer from the misleading effects of that earlier "learning".

#### Research on Corrective Advertising

The crucial questions, of course, are, "Does corrective advertising do what it is intended to do, counteract previous deception? Does it accelerate unlearning of the original response or accelerate decay of the advertising effect? If so, does the corrective information about a given brand generalize a negative image to the product class, the image of the advertiser or to the credibility of the retailer?" Only two sets of studies have examined these issues. The first is a series of studies by Keith Hunt (1972a, 1972b, 1973). Hunt used the allegedly deceptive ad originally published by Standard Oil of California touting its miraculous new additive F-310. One group saw only the original ad. Two groups saw a counter advertisement, one supposedly produced by a competitor and the other a public service advertisement from a consumer organization. Two additional groups saw a corrective ad, one supposedly voluntarily produced by Chevron and the other apparently published by Chevron under orders of the Federal Trade Commission. His results indicated that exposure to a corrective or counter ad definitely had a negative influence on the favorableness of attitudes toward the product. Further, he found no source effect. Whether the "correction" came from Standard Oil itself, a competitor, a government agency or an independent consumer organization, the results were identical - in the direction of neutralizing the positive attitudes created by the original advertisement. Further, it made little difference whether the second ad made an explicit point-by-point

attack on each deceptive statement or merely a general unsupported statement such as "previous advertising by Chevron F-310 was found to be false and deceptive." Only the ad by a competitor without explicit support did not significantly change attitudes in a negative direction.

Hunt then asked, "What happens to consumers' beliefs about the sincerity, honesty and expertise of the errant company when a corrective ad is ordered?" His data indicated that an explicit attack reduced the overall credibility of other advertising for the same brand. Further, seeing a corrective or counter advertisement hurt the reputation of the company. On the variables of honesty and sincerity, the attitude scores dropped significantly while on an expertness dimension the scores did not change. Apparently, a deceptive ad is perceived as a result of dishonesty and insincerity and not a mistake due to lack of expertise. It seems a corrective ad is a liability to the perceived honesty of the firm and not just the particular product being deceptively advertised. Interestingly, the negative attitudes toward the Chevron brand did not generalize to other gasoline brands.

A basic fear of advertisers is that the corrective advertising attacks might generalize and reduce the credibility of all advertising in general. Based on Hunt's study, the fear does not seem to be well founded. The study indicated that the negative attitudes toward Chevron F-310 did not generalize to "advertising in general" or to other brands and other product classes. He did find, however, that the honesty, sincerity and expertness of the Federal Trade Commission was considerably enhanced when subjects saw a corrective ad mentioning the name of the regulatory agency.

The second set of studies were conducted by Dyer and Kuehl (1973, 1974). They studied not only print advertising but radio ads as well. The products selected were a diet soft drink and a suntan lotion, with two sources of information for the correcting message. One of these was the company itself with no mention of a regulatory agency, and the other was presented as a Consumer Bulletin of the Federal Trade Commission. The results indicated that in print, but not broadcast advertising, "intentions to buy" decreased when the corrective advertisement was identified as an FTC source, but did not decrease when the source of the advertisement was identified as company originated. Further, under certain conditions, when the FTC was identified as the source of the message, the company was perceived as being more unscrupulous as compared to a control group. When the company was identified as the source of the message, the firm was judged to be more trustworthy. In general, however, the results support those of Hunt. Corrective advertising is effective in undoing the damage of a deceptive ad, both in attitudes and "intentions to buy". Testimony from ITT Continental Baking Co. officials presented in adversary arguments to the Commission further makes the claim that Profile Bread corrective ads caused a drop in sales of from 20-25%. If this claim is accurate, it presents further evidence that corrective advertising not only affects learned attitudes toward the brand and "intentions to buy", but produces an actual drop in sales.

#### Purpose of This Study

The purpose of this study was to contribute a third study to the precious few that exist on the effects of corrective advertising. Further, whereas the Hunt and Dyer and Kuehl studies examined the effect of corrective advertising on the image of the manufacturer in addition to the effects on the advertised product, they did not examine the effects of corrective

advertising on the retailer carrying the product. The present study used a local newspaper advertisement for a brand of product not nationally recognized and attempted to assess the impact of corrective advertising on the local retailer required to place the advertisement above his own name. Because of their particular experimental designs, both the Dyer and Kuehl and the Hunt studies presented subjects with deceptive and corrective ads out of context. In the Hunt study, subjects were presented with the ads in a booklet of questionnaires. Dyer and Kuehl presented subjects with a series of ads among which were the experimental stimuli. This study presented subjects with the newspaper in which the critical stimuli were placed.

### Methodology

The product chosen for the study was a safety device used by some motorcyclists. Both for reasons of convenience and to obtain a homogeneous population with some awareness of the existence and use of the product, the subjects of the study were male and female undergraduate college students at the University of Maryland and at George Washington University.

The basic design of the study was as follows:

The experimenters approached students in the classroom and presented them with photostatic copies of five pages of the June 13, 1973, issue of the New York Daily News. They were told that this was a study of newspaper readership and were asked to please read the paper in the same manner that they usually read any newspaper. They were asked not to study or memorize the stories and advertisements, but rather go through the paper as they usually might, skimming the stories of little interest and perhaps spending more time with those of greater interest. They were given about 10 minutes to do so.

Within the newspaper one of the published ads had been clipped out and the original deceptive safety device ad inserted before photostating, so each paper in this phase included an ad for the device supposedly sold in New York by a retailer which could well have existed in New York. The ad claimed to offer protection with the world's safest device: one piece for lighter weight with increased strength. Government approved.

After reading the newspaper, including the deceptive ad, subjects were handed a questionnaire and were asked to:

1. Remember the stories they read or noticed and identify each on an open end questionnaire.
2. Complete a multiple choice "test" on several of the stories.
3. Identify each ad they remembered reading or seeing (unaided).
4. Check the ads they noticed from a list. (Some of the ads, in fact, did not exist in the newspaper and others did. This was a measure of aided recall.)
5. Rate another advertised product (bathing suits) on a semantic differential attitude scale.
6. Rate the retailer of the other product.
7. Rate the safety device and the retailer of these devices.



The purpose of asking irrelevant questions about newspaper stories and other ads was to cover the real purpose of this study, the ads for safety devices.

Upon completing this phase of the study, the questionnaires were collected and the subjects presented with some pages of the June 18, 1973 edition of the New York Daily News. In this version, an actual ad in the paper was replaced with one of eight versions of the corrective ads to be studied. Each corrective advertisement contained at least the following copy:

The advertisement shown above appeared in this newspaper. We have now been informed that the (device) advertised is not the world's safest (device). The (device) in question is made of (plastic). Although such a (device) may meet (government) standards, no (plastic device) has yet met the most rigorous standard set out under the test procedures to qualify for (agency) approval.

Because of the error under which this device was advertised and sold, (Retailer) will, for 30 days following the appearance of this notice, accept the return for full cash refund on any (device) purchased as a result of the above advertisement.

Sixty-four subjects were then asked to read this newspaper as they had the earlier one and then complete a second questionnaire. This questionnaire again included questions on advertisements in the newspaper and repeated the semantic differential scales on the safety device found in the first phase questionnaire.

Upon completion the subjects were told that the study concerned advertising and were asked to go back to the newspaper and carefully reread the ad for the safety device. This phase of the study is similar to the work of Hunt and of Dyer and Kuehl in that subjects specifically focused on the corrective ad, rather than having it buried in the newspaper.

At this point, the subjects were again asked to complete the semantic differential scale on the product. After collecting several other minor questions and some demographic information, the subjects were debriefed and told the real purpose of the study. All questions were answered to alleviate any concern that may have been generated in the subjects since, to a small extent, they had earlier been deceived about the true purpose of the experiment.

The experimental methodology consisted of a 2 x 2 x 2 x 15 analysis of variance design. The variables consisted of three copy and layout variations and the 15 item semantic differential scale. The 64 subjects were randomly assigned, 8 subjects to each of the eight cells. Each experimental treatment or advertising variation was identical in the placement in the newspaper, in copy, in graphics, and in text, except for the experimental manipulations.

A second group of 32 subjects (control group) were presented with only the June 18 edition of the newspaper, each containing one of the eight versions of the corrective ad. They were not exposed to the first phase of the study and did not see the original deceptive advertisement. A comparison of the control group with the experimental group would indicate the effect on consumers who are exposed to a corrective ad, but who did not see the

original deceptive ad. Further, the data may be interpreted as a measure of the biases inherent in the experimental design. Although every effort was made to cover the purpose of the study, it may be that the experimental subjects were more sensitized to observing the second ad, since this group had been previously exposed to the same product and a similar advertisement.

### Results

The first analysis considered the effect of the eight variations of advertising stimuli. Among the 64 experimental subjects exposed to both the original ad and one of the corrective versions in the natural context of a newspaper, the data indicated no significant three-way interaction effect (repeated measures analysis of variance design  $F=1.046$ ; 1/56 df) and no significant two-way interaction effects. Further, none of the copy and layout variables (main effects) were significantly different from each other. Results of analyses of variance for the 32 control subjects were similar. In short, stimulus variations in copy and layout had no significant effect in this study. Since the text and layout made no difference in the effectiveness of the corrective message, the data were pooled into a single experimental and a single control group.

#### Effect of Corrective Advertising on Attitude Change

Using the semantic differential scale as a measure of product perception, the effect of being exposed to a deceptive advertisement and then a corrective advertisement can be seen in Figure 1. The broken line represents attitude scores of the 64 subjects after having seen only the original deceptive ad and before exposure to the corrective ad. The solid line represents semantic differential scores after having been exposed to the corrective ad. Upon having been exposed to the corrective ad, the product is perceived as being less reliable, less effective, less believable, more worthless, inferior, deceptive, mistaken, misleading, more poorly made, and less recommendable.

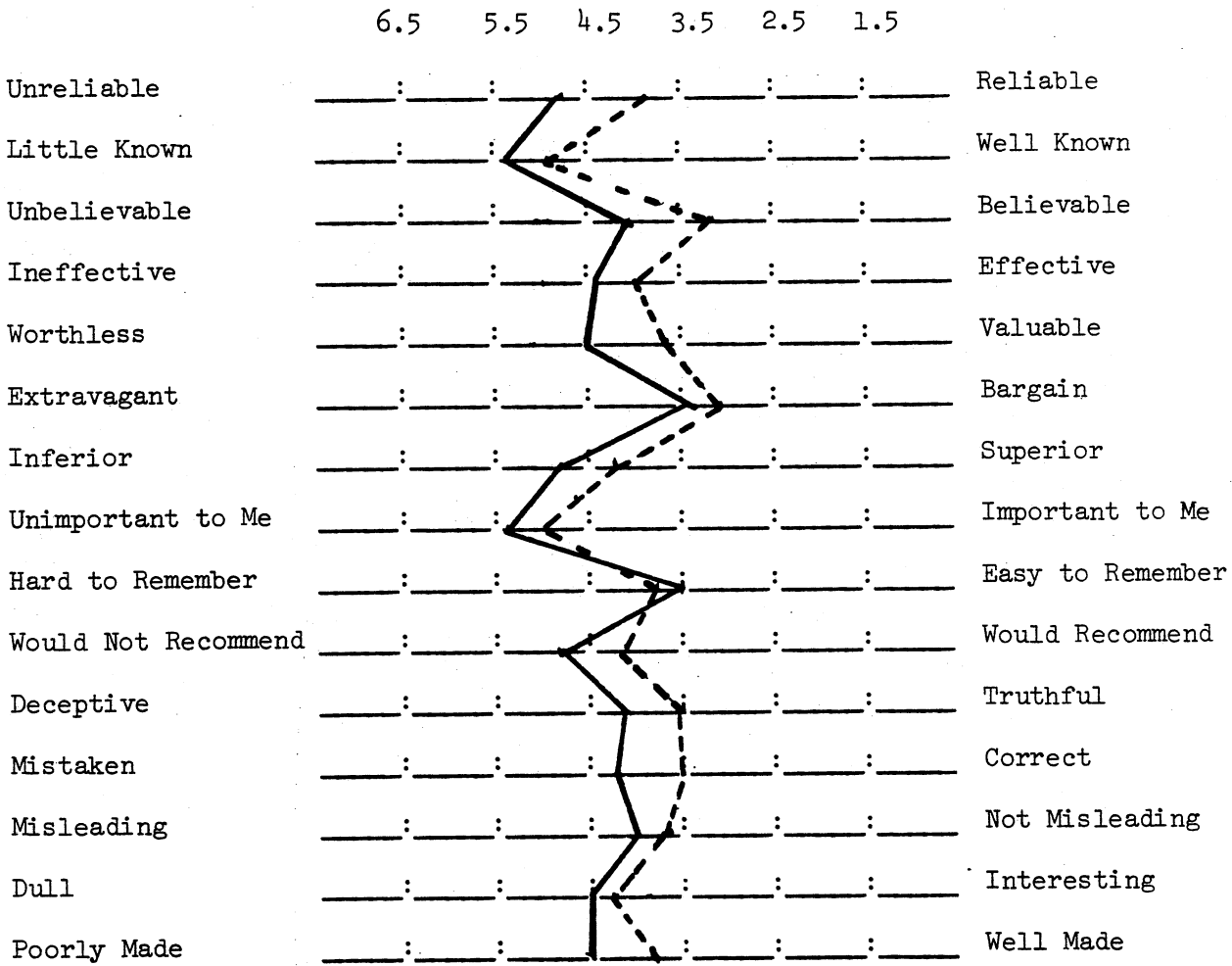
The two profiles are significantly different from each other at the .01 level of confidence (repeated measures design analysis of variance  $F = 7.54$ ; 1/56 degrees of freedom). Interpretation of the results indicates that the overall effect of presenting a corrective ad changes perceptions of the item advertised as measured by the semantic differential in a direction that makes it less desirable; the change is in a negative direction.

#### Perception of the Retailer

The purpose of these particular ads was to change views of the product and not the retailer. The question that arises, however, is, "Having seen a corrective ad for a product that is now perceived as more negative, do the subjects generalize the negative view not only to the product but also perceive the retailer as being more unreliable and more negative?"

Figure 2 presents these results. The differences in attitudes before and after having seen the corrective ads are not statistically significant. The overall profiles do not differ from each other, with the variations likely due to chance.

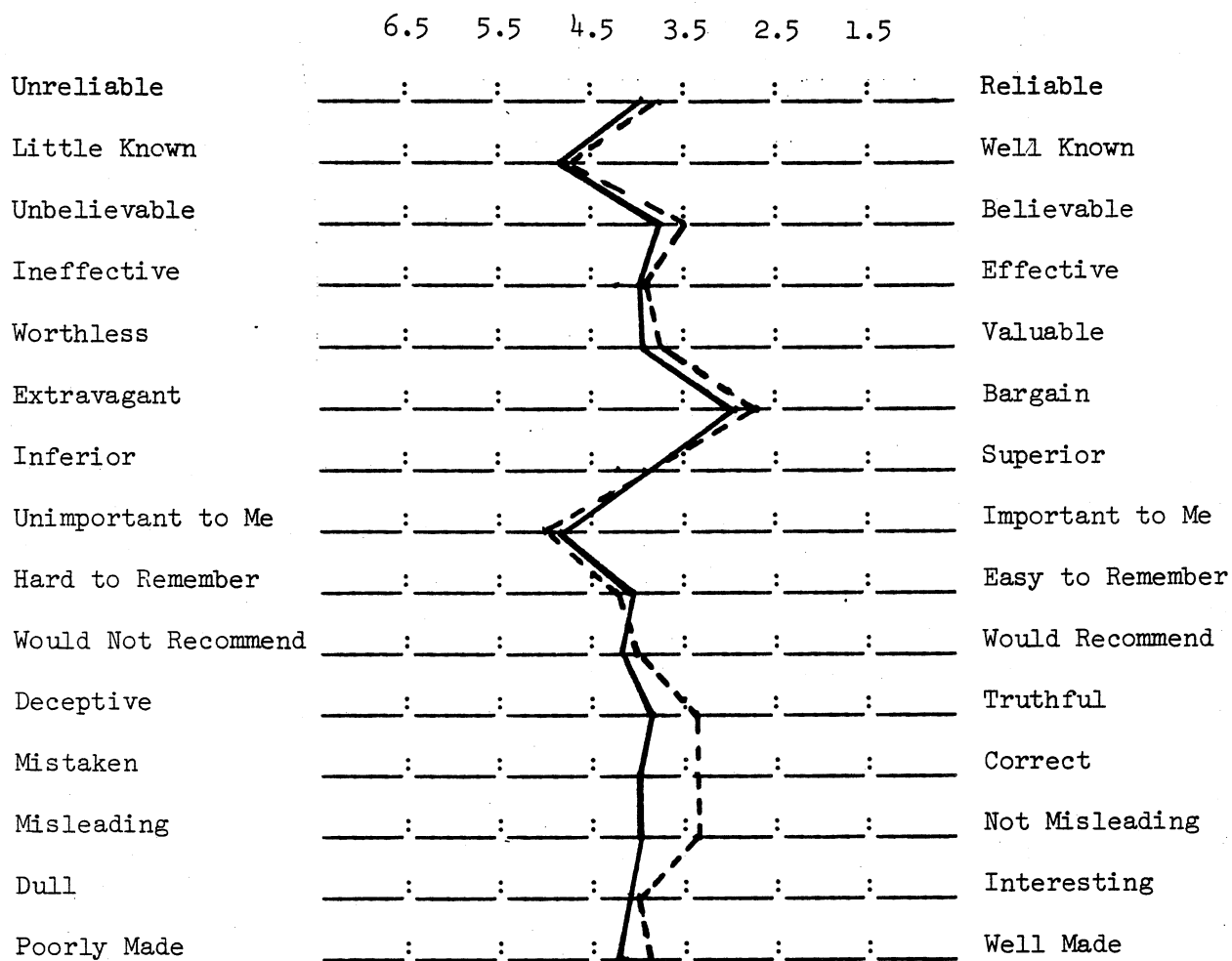
Both Hunt and Dyer and Kuehl had found that under certain conditions the manufacturer of a product with a corrective ad was perceived as less



----- Results after seeing original ad  
 \_\_\_\_\_ Results after seeing both original and corrective ad

(NOTE: On this figure, not in statistical analysis, missing data are ignored. In the questionnaire the positive and negative poles of the adjectives were randomized.)

Figure 1. Semantic differential scales (safety device).



----- Results after seeing original ad  
 \_\_\_\_\_ Results after seeing both original and corrective ad

(NOTE: On this figure, not statistical analysis, missing data are ignored. In the questionnaire the positive and negative poles of the adjectives were randomized.)

Figure 2. Semantic differential scales (retail store)

honest and sincere, hence it would be reasonable to expect that if a retailer published an ad for a product that is perceived as deceptive, the retailer himself might also be perceived as deceptive, mistaken and misleading. Nevertheless, the data in this study indicate that the retailer is not affected by publishing a corrective advertisement. Whereas attitudes toward the product become significantly more negative, the attitudes toward the retailer do not change significantly.

#### Effects of Forced Exposure

As already mentioned above, the previous studies in corrective advertising presented the ads out of context, under forced exposure conditions. In this study, the ads were presented in the context of a newspaper. However, upon completion of the first two phases of the study, subjects were specifically asked to turn to the corrective ad, study it carefully, and then again asked to complete the semantic differential measure. The results indicated no significant difference in profiles toward the product having viewed the ad under "natural" conditions and under conditions of forced exposure. For example (and simplicity), the overall mean of all subjects over all scales for the forced exposure conditions was 4.78, and under the non-forced condition 4.52. The difference is not significant at the .05 level of confidence. However, under conditions of forced exposure, one of the copy and layout variables suddenly took on some significance. Apparently under natural "reading" conditions minor changes in the message itself make little difference, but upon forced exposure, changes in the text and layout do take on significance.<sup>2</sup> Again, these findings were quite similar to the previous studies. On an overall basis, it makes little difference whether subjects are shown the ad under natural conditions or on a forced exposure condition.

#### Control Conditions

Whenever a corrective advertisement is published, two groups of readers will be exposed to the correction. One group are those who saw the original deceptive advertisement, and in the other group are those who did not see the original ad and are being exposed to the product advertisement only under the corrective conditions.

To examine what the effects would be on readers who did not see the original advertisement, 32 control subjects were exposed to one of the versions of the corrective ads. Under both forced exposure and non-forced exposure conditions, none of the layout or copy variables nor interaction effects were significant.

The profile of the semantic differential scales for rating of the product for the control group (did not see the original ad) under non-forced exposure conditions (overall mean = 4.521) was not significantly different from the mean of the experimental group who saw both ads (overall mean: 4.520). Under forced exposure conditions, the overall mean of the profile toward the product of the experimental group who saw both the original and the corrective ad was 4.78. For the control group who saw only the corrective ad, the overall mean was 5.06. Again, neither these means nor the profiles differ significantly. These findings imply that exposure to a corrective ad produces perceptions toward the product that are quite similar regardless of whether or not the consumer has been exposed to the original deceptive ad.

In summary, the results of this study indicate the following findings:

1. Reading a corrective ad in a newspaper tends to undo some of the positive attitudes toward a product created by a deceptive advertisement.
2. Negative effects created by a corrective advertisement do not generalize to the local retailer who placed the ad, in this case for an item not identified by a nationally known brand name.
3. When exposure to an ad is forced upon a subject, the profile changes that result are not significantly different than if he is exposed to it under more natural conditions.
4. Subjects who did not see the original ad but were exposed only to the corrective ad had identical profiles toward the product as did subjects who saw both ads.

The results of this study support the findings of previous work that exposure to a corrective advertisement does tend to undo some of the attitudes created by misleading, deceptive and unfair advertising at least in laboratory settings. The ITT claims of a 25% drop in sales of Profile Bread adds further data. Unfortunately, just as the research supporting the placement of corrective advertising is beginning to emerge, the Federal Trade Commission appears to be turning away from the concept and returning to merely requiring "cease and desist" action. Every single corrective advertising order prepared by the staff and contested by the respondent has been turned down by either the adjudicative judge or by the Commission on appeal. It is unfortunate, for as the Washington Star newspaper has editorially pointed out, ". . . the sponsor who paid to mislead the public must now pay again to correct the previous deception is the essence of (corrective advertising). A diabolically clever scheme and a simple one that makes monumental good sense" (Randal, 1973).

#### FOOTNOTES

1. The contributions of Jay S. Haladay, Masao Nakanishi and Richard Lutz in the data processing phase of this study are gratefully acknowledged. Philip G. Kuehl and Robert F. Dyer were instrumental in arranging the availability of subjects. H. Keith Hunt's critical advice is deeply appreciated. Portions of this study were carried out while the authors were associated with the Federal Trade Commission. This paper represents the interpretations and opinions of the authors, and not those of the Federal Trade Commission or any individual Commissioner.
2. This conclusion must be somewhat tempered by the fact that because of experimental exigencies we did not use a second control group that viewed the ads only under forced conditions. The forced condition subjects were the same ones that had previously been exposed to the corrective ad under non-forced conditions.

#### REFERENCES

- Anon. Corrective advertising and the FTC: No, Virginia, Wonder Bread doesn't help build strong bodies twelve ways. Michigan Law Review, 1971, 70, 374-399.

- Anon. Corrective advertising--the new response to consumer deception. Columbia Law Review, 1972, 72, 415-431.
- Capon, N. & Hulbert, J. The sleeper effect: An awakening. Public Opinion Quarterly, 1973, 37, 333-358.
- Dyer, R. F. & Kuehl, P. G. Source and strength effects in corrective advertising. In S. Ward & P. Wright (Eds.), Advances in consumer research. Vol. 1. Urbana: Association for Consumer Research 1973. Pp. 85-86 (abstract).
- Dyer, R. F. & Kuehl, P. G. The 'corrective advertising' remedy of the FTC: An experimental evaluation. Journal of Marketing, 1974, 38 (1), 48-54.
- Hovland, C. I., Lumsdaine, A. A. & Sheffield, F. D. Experiments in mass communication. New York: John Wiley and Sons, 1949.
- Hunt, H. K. Reception, inoculation, attack: Implications for inoculation theory, public policy, and advertising strategy. Unpublished doctoral dissertation, Northwestern University, 1972.
- Hunt, H. K. Source effects, message effects, and general effects in counter-advertising. In M. Venkatesan (Ed.), Proceedings, third annual conference, Association for Consumer Research, 1972.
- Hunt, H. K. Measuring the impact and effectiveness of counter messages. Paper presented at the American Marketing Association Conference, "Advertising and the Public Interest," Washington, D.C., May, 1973.
- Kassarjian, H. H. Applications of consumer behavior to the field of advertising. Journal of Advertising, 1974, 3, 10-15.
- Randal, J. Washington close up: 'Unlearning' the ad effects. Evening Star and Washington Daily News, Thursday, June 14, 1973, Page A 20.
- Ray, M. L., Sawyer, A. G. & Strong, E. C. Frequency effects revisited. Journal of Advertising Research, 1971, 2 (1), 14-20

THE RELATIONSHIPS AMONG ADVERTISING, ECONOMIC DEVELOPMENT  
AND CONSUMER WELFARE: AN EXAMINATION OF RECENT  
WORLD-WIDE EVIDENCE

Anthony F. McGann and Nils-Erik Aaby<sup>1</sup>  
University of Wyoming

In the research reported here, the relationships among advertising, economic development and consumer welfare are examined using data from 47 countries in the period 1970-1972. Aggregate advertising was found to be significantly associated with gross national product, with foreign trade and with per capita GNP in the overall group of countries.

Attention was then focused on the relationships among these variables in the less-developed countries of the sample. In this group, advertising was found to be strongly associated with aggregate indices of economic development, but virtually uncorrelated with variations in per capita GNP. The authors infer that the uncontrolled growth of population in the less-developed countries prevents advertising from playing its usual role as the institution which distributes wealth to individual consumers.

#### Introduction

In 1958, Drucker noted a global phenomenon which he regarded as a new and major threat to international tranquility, namely, the growing discrepancy between the economic well-being in industrialized and "underdeveloped" countries. More recently, Ward (1968) has advanced this same thesis. Economists prefer to discuss the problem of this discrepancy under the name "terms of trade" (e.g., Kindleberger, 1956). By either name, the investigators of this phenomenon can be divided into two camps; those who are pessimistic about the prospects for closing, or at least diminishing, the gap and those who are optimistic over solutions to this problem.

An example of the pessimistic view is expressed by Prebisch (1966) who argues that the underdeveloped countries, essentially selling commodities in extremely price-competitive markets, are at a long-term disadvantage relative to the industrialized nations which trade differentiated goods in markets characterized by less-intense price competition. In contrast, the optimists argue that global discrepancies in economic well-being can be reduced. Recent experience has shown that underdeveloped nations selling energy commodities can achieve dramatic changes in foreign trade balances by embargoes, quotas and concerted price increases. More fundamentally, increased marketing activity by underdeveloped nations offers these countries the prospect of a better "deal" in world marketplaces and subsequent improvements in the economic well-being of their populations. Drucker ascribes to marketing this significant role:



. . . in every 'under-developed' country I know of, marketing is the most underdeveloped--or least developed--part of the economy . . . these countries are stunted by inability to make effective use of the little they have. Marketing might by itself go far toward changing the economic tone of the existing system--without any change in methods of production, distribution of population, or of income. (Drucker, 1958).

This study was conceived as a partial test of the optimistic view toward the solution of international economic disparities. The authors adopt the position that recent benefits to underdeveloped nations which result from restrictions in trade (e.g., embargoes and quotas) are of a short-run nature. Alternatively, the benefits from intensified marketing activity hold the promise of more permanent improvements to economic progress and social welfare in the underdeveloped areas of the world. Specifically, this study examines the apparent relationships which have existed in recent years between advertising and national economic development, and between advertising and individual consumer welfare.

It is recognized, of course, that advertising is only one aspect of marketing activity. In this respect, the present study is not a test of the effect of overall marketing activity. Nevertheless, advertising has three properties which make it an intellectually-attractive point from which to begin the assessment of contemporary marketing effects:

1. Functionally, advertising is generally recognized as a stimulant to demand. Thus, any nation engaged in international trade, but particularly the underdeveloped countries would benefit from the effects of advertising.<sup>2</sup>
2. Operationally, advertising expenditures are measurable and available. In contrast, other aspects of international marketing, e.g. relative strengths of distribution channel members, represent, at best, qualitative data.
3. Philosophically, advertising has been described in domestic terms, as the institution of abundance; more precisely, the institution by which abundance is broadly distributed among populations of the industrialized western world (Sandage, 1972). If advertising can distribute "abundance" in the industrialized areas of the world, perhaps it can be expected to contribute, in the underdeveloped areas of the world, to the economic progress of consumers, or at least to their survival.

#### Design of the Study

In order to assess the relationships among advertising, economic development and individual consumer welfare, data from two sources were collected. First, conventional indices of the national economy (Gross National Product, Exports, Imports and Population) were obtained from the International Monetary Fund (International Financial Statistics, 1974), for the period 1970-1972. This period was chosen since it represents the most recent three-year period for which complete data were commonly available. Second, by-country advertising expenditures ("Advertising Sales") were collected from annual special editions of Advertising Age (1973, 1972, 1971).

After excluding those countries for which complete data were not available for at least two of the three years under study, a group of 47 countries were available for investigation. Using IMF designations, about 28% of this group can be considered "industrialized;" almost 22% are characterized as "other developed areas;" the remaining half are described as "less developed areas." (See Appendix 1 for country list and designations).

All financial data were converted to U.S. dollars using contemporaneous exchange rates for that year. The United States was intentionally excluded from this analysis; many of the centrally-directed economies and some of the newest nations had insufficient data to be included. For the countries included in the study, two substantive hypotheses were proposed:

- HR<sub>1</sub>: Across nations, variations in economic activity are positively and significantly related to variations in advertising expenditures.
- HR<sub>2</sub>: Across nations, variations in individual consumer welfare are positively and significantly related to variations in advertising expenditures.

#### World Economic Conditions During the Period

During the period under study, economic conditions in the sample countries can be characterized as expansive. Table 1 shows consistent increases in mean values for GNP, Exports and Imports. It can also be seen that mean values for aggregate advertising and for "advertising per employee"<sup>3</sup> have increased each year. However, the substantial variations about these mean values suggests a more disaggregated analysis.

TABLE 1

#### Overview of International Economic Activity; 1970-1972

Index	Mean ( $\bar{X}_n$ ) (Million Dollars)	St. Dev. ( $\sigma_n$ )
1970:		
Gross National Product	28,182.	52,177.
Exports	4,508.	7,315.
Imports	4,845.	7,006.
Advertising "Sales"	128.	288.
Advertising "Sales" per employee	.058	.039
1971:		
Gross National Product	30,849.	53,942.
Exports	4,929.	4,963.
Imports	5,212.	7,522.
Advertising "Sales"	133.	291.
Advertising "Sales" per employee	.063	.360
1972:		
Gross National Product	33,774.	67,015.
Exports	6,039.	9,733.
Imports	6,115.	9,033.
Advertising "Sales"	150.	332.
Advertising "Sales" per employee	.069	.041

It has already been suggested that the level of industrialization within a country is associated with its economic development. Table 2 shows the results of different levels of industrialization on both aggregate and per capita GNP. In the industrialized countries both aggregate and per capita GNP have increased each year. For the remaining countries, both figures increased slightly from 1970-71 and then decreased in the subsequent year. Table 2 also details the effects of different levels of industrialization on foreign trade as well as differences in the advertising industry across groups of countries. In the advertising industry, for both gross billings and billings per agency employee, the major differences exist between the industrialized countries and all other countries in the sample. For national product, foreign trade and the advertising industry, the between-group differences in economic activity are highly significant.

A second method of reducing heterogeneity among the sample countries is to form subsamples of countries on the basis of their common membership in economic unions. Approximately half of the countries studied belonged to one of four major unions. Table 3 presents differences for these "member" countries on aggregate and per capita GNP, foreign trade and the advertising industry respectively. Again, there are generally significant between-group differences.

By either method of disaggregation, the variances noted about mean values in Table 1 are translated into stark differences between countries which presently enjoy the benefits of past industrialization and those which are not (yet) industrialized. Using the IMF country categories, sharper differences are revealed between the "have" and the "have not" nations than are shown by use of membership in economic unions. It is clear that the discrepancies Drucker (1958) hoped marketing activity would alleviate have not disappeared in the ensuing 15 years.

#### General Relationships Among Advertising, Economic Development and Individual Consumer Welfare

For the period under study one measure of advertising, aggregate billings, was examined for its relationship to national product and foreign trade. Table 4 shows a summary of the results when gross billings are used, in bivariate regression, to explain variation in national product and in foreign trade for the entire sample of 47 countries. These regression results suggest that aggregate advertising billings vary directly, positively and significantly with both economic activity and with individual consumer welfare. Further, the variation explained in these "dependent" variables ranges from a high of better than 80% for GNP to a low of about 20% for per capita GNP; better than 60% of the by-country variance in foreign trade (exports and imports) is accounted for by variance in the by-country aggregate advertising expenditures. Finally, the authors are inclined to note that the almost perfect correspondence between national sales and national advertising is no more a guarantee of maximized GNP than it is a guarantee of a firm maximizing sales in the analogous micro setting.

#### Relationships Among Advertising Economic Development and Individual Consumer Welfare in "Less Developed" Areas

This study has taken as its focus the economic conditions in the under-developed nations. Accordingly the role of advertising in the twenty-four countries identified by IMF as "less-developed" was analyzed in regression

TABLE 2  
Effects of Industrialization<sup>a</sup> on:

Year Group <sup>b</sup>	National Product			Foreign Trade			Advertising Industry			
	Mean (Billion Dollars)	df	F- ratio <sup>c</sup>	Mean (Billion Dollars)	df	F- ratio <sup>c</sup>	Mean (Million Dollars)	df	F- ratio <sup>c</sup>	
1970	GNP	79.9		Exports	13.0		Advertising	394.2		
		12.6			1.7		Sales	47.4		
		5.5	2,44	14.8		2,44		11.8	2,43	
1971	GNP	86.9		Exports	14.2		Advertising	411.4		
		14.3			1.9		Sales	50.2		
		6.1	2,44	17.4		2,44		11.6	2,43	
1972	GNP	102.5		Exports	17.3		Advertising	460.9		
		11.7			2.4		Sales	57.6		
		4.4	2,44	16.5		2,44		13.3	2,44	
1970	Per Capita GNP	(Dollars) 2,828 1,190 413		Imports	13.4 2.6 0.9		Advertising Sales per Employee	(Thousand Dollars) 95.9 41.6 42.4		
			2,44	74.4		2,44			2,44	13.2
1971	GNP	3,254 1,343 424		Imports	14.3 2.9 1.0		Advertising Sales per Employee	102.7 55.5 42.0		
			2,44	79.5		2,44			2,44	22.7
1972	GNP	3,695 847 278		Imports	17.0 3.2 1.1		Advertising Sales per Employee	118.7 54.7 46.1		
			2,44	88.9		2,44			2,44	30.5

<sup>a</sup>Degree of Industrialization taken from International Financial Statistics (1974).  
<sup>b</sup>1 = "Industrialized Countries;" 2 = "Other Developed Areas;" 3 = "Less Developed Areas"  
<sup>c</sup>All F-ratios significant at  $p < .001$  (one-way ANOVA)

TABLE 3  
Effects of Membership in Trade Associations on:

Year	Group <sup>a</sup>	National Product			Foreign Trade			Advertising Industry		
		Mean (Billion Dollars)	df	F- ratio <sup>b</sup>	Mean (Billion Dollars)	df	F- ratio <sup>b</sup>	Mean (Million Dollars)	df	F- ratio <sup>b</sup>
1970	1	GNP	97.2		Exports	17.7		Advertising	322.2	
	2		29.1			5.4		Sales	198.6	
	3		9.0			1.1			32.8	
	4		0.9	3,20		0.2	3,20		2.0	3,19
1971	1	GNP	116.4		Exports	20.3		Advertising	365.8	
	2		34.4			6.0		Sales	213.1	
	3		10.2			1.2			29.6	
	4		0.9	3,20		0.2	3,20		1.6	3,19
1972	1	GNP	130.9		Exports	24.9		Advertising	417.7	
	2		35.9			7.0		Sales	206.2	
	3		7.2			1.4			33.0	
	4		0.7	3,20		0.2	3,20		1.4	3,20
1970	1	Per Capita GNP	(Dollars) 2,576		Imports	18.8		Advertising	(Thousand Dollars) 95.5	
	2		2,556			6.4		Sales per Employee	62.4	
	3		543			1.0			51.7	
	4		387	3,20		0.2	3,20		39.5	3,20
1971	1	GNP	3,083		Imports	19.9		Advertising	114.1	
	2		2,992			6.9		Sales per Employee	76.9	
	3		556			1.2			38.3	
	4		406	3,20		0.2	3,20		44.8	3,20
1972	1	GNP	3,457		Imports	23.9		Advertising	129.3	
	2		3,240			8.0		Sales per Employee	84.6	
	3		155			1.4			43.7	
	4		305	3,20		0.3	3,20		43.6	3,20

<sup>a</sup>1 = European Common Market; 2 = EFTA; 3 = CAEC; 4 = LAFTA

<sup>b</sup>F-ratios significant at  $p < .05$  (one-way ANOVA)

<sup>c</sup>not significant at  $p < .05$

similar to that for the total country group. Table 5 shows the summary of the bivariate regression results when aggregate advertising is used to explain variance in national product as well as in foreign trade.

TABLE 4  
Bivariate Regression; Advertising on Gross National  
Product and Foreign Trade

(N = 47)				
<u>Dependent Variable</u>	<u>Intercept</u>	<u>Normalized β</u>	<u>R<sup>2</sup></u>	<u>F-ratio<sup>a</sup></u>
GNP (1970)	6.553	.934	.873	302.273
GNP (1971)	8.649	.901	.813	190.770
GNP (1972)	5.597	.934	.872	299.488
Exports (1970)	1.805	.833	.694	99.598
Exports (1971)	1.979	.811	.658	84.839
Exports (1972)	2.452	.818	.670	89.217
Imports (1970)	2.382	.792	.628	74.144
Imports (1971)	2.385	.823	.677	92.352
Imports (1972)	2.861	.800	.640	78.175
Per capita GNP (1970)	1015.671	.483	.233	13.362
Per capita GNP (1971)	1141.615	.455	.207	11.463
Per capita GNP (1972)	997.918	.490	.241	13.934

<sup>a</sup>F-ratios significant at  $p < .005$ .

In this portion of the analysis, the results are mixed. For the less-developed countries, aggregate advertising is positively, directly and significantly associated with GNP, exports and imports. For this subsample of countries, aggregate advertising explains about the same amount of variance in GNP and import values as it does for the overall sample; somewhat less variance in exports is explained than was the case for the overall sample. However, aggregate advertising values are not significantly related to variation in per capita GNP. Thus, while advertising bears a common relationship to GNP and foreign trade in both the overall sample and in the less-developed countries, it is not influential in accounting for differences in individual consumer welfare for the latter group.

For the subsample of "less developed" countries, a second measure of advertising was developed. This measure is the ratio of advertising to gross national product. This measure differs from the absolute advertising level used above and represents the proportion of national product value which is devoted to advertising.

Again regression was employed to test the relationship between this proportional advertising figure and per capita GNP. Table 6 contains the regression results for the three years studied. As before, there is virtually

no relationship between advertising (relative to GNP) and per capita GNP.  
(See Table 7 for a summary of results relative to the research hypotheses.)

TABLE 5

Bivariate Regression: Advertising on Gross National Product and  
Foreign Trade in the "Less Developed" Countries

(N = 24)				
<u>Dependent Variable</u>	<u>Intercept</u>	<u>Normalized β</u>	<u>R<sup>2</sup></u>	<u>F-ratio</u>
GNP (1970)	2.849	.913	.833	104.891*
GNP (1971)	2.255	.916	.840	110.327*
GNP (1972)	1.006	.921	.848	116.846*
Exports (1970)	.639	.704	.495	20.613*
Exports (1971)	.646	.681	.463	18.140*
Exports (1972)	.906	.638	.407	14.401*
Imports (1970)	.723	.699	.489	20.124*
Imports (1971)	.691	.783	.613	33.298*
Imports (1972)	.851	.787	.619	34.100*
Per capita GNP (1970)	403.418	.099	.009	.210**
Per capita GNP (1971)	382.745	.225	.050	1.115**
Per capita GNP (1972)	262.590	.079	.006	.114**

\*F-ratio significant at  $p < .001$

\*\*F-ratio not significant at  $p \leq .20$

TABLE 6

Bivariate Regression: Dollars of Advertising Per Dollar of  
GNP on Per Capita GNP in the "Less Developed" Countries

(N = 24)				
<u>Dependent Variable</u>	<u>Intercept</u>	<u>Normalized β</u>	<u>R<sup>2</sup></u>	<u>F-ratio</u>
Per capita GNP (1970)	1501.383	.184	.033	.768**
Per capita GNP (1971)	1137.236	.327	.106	2.612*
Per capita GNP (1972)	946.848	.240	.058	1.349**

\*Significant;  $p < .10$

\*\*Not significant;  $p \leq .20$

TABLE 7  
Hypotheses and Analytic Results

	All Countries Studied	Less-Developed Countries Studied
HR <sub>1</sub> : Across nations, variations in economic activity are positively and significantly related to variations in advertising expenditures	SUPPORTED	SUPPORTED
HR <sub>2</sub> : Across nations, variations in individual consumer welfare are positively and significantly related to variations in advertising expenditures	SUPPORTED	NOT SUPPORTED

#### Limitations and Conclusions

The authors acknowledge several limitations to this study. First, about half the nations of the world were not available for study because they did not have sufficiently complete data. Notably-frequent exclusions for this reason occur in the Sino-Soviet countries and among the newest nations. Second, production and population figures are less reliable for some countries than for others. However, for purposes of this study, the published data from two independent, reputable sources were accepted at face value. Third, no attempt has been made to assess the effect of advertising in a given period with national product or foreign trade in later periods. Finally the authors elected to sacrifice the trend-fitting possibilities (of including earlier periods) in order to isolate any uniqueness which may have been a feature of the most recent international economic experience.

In spite of these limitations, it can be asserted that advertising bears a strong correlative relationship to the national economic condition in the countries studied. Yet, while advertising explains significant levels of variation in GNP and in foreign trade in the less-developed nations, its explanatory power with respect to individual consumer welfare is virtually nil. The discrepancy between good explanations of GNP and no explanation for per capita GNP must inferentially be attributed to population growth.

The authors acknowledge one more item of evidence on the effect of population to individual consumer welfare. We recognize that growth in per capita GNP does not guarantee equitable distribution of incomes across a national population; rather it merely serves as an economic prerequisite for the political decisions which would accomplish this distribution. Finally we conclude that in our partial test of Drucker's proposal, advertising of itself has not been significantly associated with individual consumer welfare in the less-developed nations.



## FOOTNOTES

1. Anthony F. McGann, Associate Professor and Nils-Erik Aaby, Graduate Assistant, Department of Business Administration, College of Commerce and Industry, University of Wyoming. The authors are grateful for the substantive comments of Professors Frank K. Reilly and Raymond A. Marquardt in the development of this manuscript. The authors also wish to thank Mary Lee Cadwell, Graduate Research Assistant, for her extensive assistance.
2. Some would argue that this benefit of stimulated demand is not available to underdeveloped countries in that they primarily sell undifferentiated commodities not amenable to advertising. The authors reject this argument for two reasons. First, it ignores advertising's potential for stimulating primary demand. Second, recent domestic experiences with salt and dressed chickens, and similar international experiences with coffee and wines suggests that some commodities are amenable to differentiation via advertising.
3. Advertising sales per employee refers to the "sales productivity" of those in the advertising industry of each country. This figure was calculated by dividing the combined billings of those agencies reporting the number of their employees by the number of employees. Billings of agencies not reporting their employment figures are excluded from this calculation.

## REFERENCES

- Drucker, P. Marketing and economic development. Journal of Marketing, 1958, 22 (3), 252-259.
- International agency section. Advertising Age, March 26, 1973, 33+.
- International agency section. Advertising Age, March 27, 1972, 29+.
- International agency section. Advertising Age, March 29, 1971, 29+.
- International financial statistics. International Monetary Fund, 27 (3), March, 1974.
- Kindleberger, C.P. The terms of trade, a European case study. New York: John Wiley & Sons, Inc., 1956.
- Prebisch, R., In Root, F. R., et al. International trade and finance. (2nd ed.) Chicago: South-Western Publishing Company, 1966 441-442.
- Sandage, C. H. Some institutional aspects of advertising. Journal of Advertising, 1972, 1 (1), 6-9.
- Ward, B. J. The lopsided world. New York: W. W. Norton and Company, Inc., 1968.

## APPENDIX I

## Country List

Australia <sup>b</sup>	Honduras	Paraguay
Austria <sup>a</sup>	Indonesia	Philippines
Belgium <sup>a</sup>	Ireland <sup>b</sup>	Portugal <sup>b</sup>
Brazil	Israel	South Africa <sup>b</sup>
Britain <sup>a</sup>	Italy <sup>a</sup>	South Korea
Canada <sup>a</sup>	Jamaica	Spain <sup>b</sup>
Chile	Japan <sup>a</sup>	Sweden <sup>a</sup>
Colombia	Kenya	Switzerland <sup>a</sup>
Costa Rica	Kuwait	Taiwan
Denmark <sup>a</sup>	Malaysia	Thailand
Dominican Republic	Morocco	Turkey <sup>b</sup>
Ecuador	Netherlands <sup>a</sup>	Uruguay
El Salvador	New Zealand <sup>b</sup>	Venezuela
Finland <sup>b</sup>	Nicaragua	West Germany <sup>a</sup>
France <sup>a</sup>	Norway <sup>a</sup>	Yugoslavia <sup>b</sup>
Greece <sup>b</sup>	Pakistan	

---

<sup>a</sup> = Industrialized countries; <sup>b</sup> = other developed areas; other countries categorized as "less-developed." (International Financial Statistics, 1974, pp. 8-9).



## INTRODUCTION: EXPERIMENTAL RESEARCH AND THEORIES OF CONSUMER CHOICE

Penny Baron and Gerald Eskin  
The University of Iowa

A principal objective in organizing this session was to bring together persons who share an interest in consumer choice but whose research starts from somewhat different perspectives.

Research on consumer choice has typically followed one of two rather distinct paths. The first of these is more common in marketing research. It is empirical in its orientation and it purports to be at least somewhat concerned with obtaining practical results. The second path is primarily theoretical. This approach derives mainly from economics and mathematical psychology. It seeks to develop theories which are rigorous and precisely formulated. The area of consumer choice has attracted a large number of highly competent researchers who pursue one or the other of these paths. But inadequate professional exchange has occurred between these two groups. Each has conducted its research uninformed by the needs and objectives of the other group.

This situation often results in research efforts which suffer from important limitations.

Much of the work based on empirical investigations of consumer choice behavior is informed by a series of conceptual frameworks rather than a single well-developed theoretical system.

Each framework suggests a lengthy list of important categories of variables. Each category in turn is composed of an extensive list of relevant variables. This procedure does result in comprehensive coverage but it is not clear how it directs empirical work. The lists give us no way to decide which variables are important and which are not. We have no basis for preferring one list over another. In the absence of an encompassing theoretical guide, it is difficult to evaluate the merits of alternative research plans in advance. As a result, the empirical "facts" discovered often turn out to be trivial, uninterpretable or relevant only to a particular time and place.

These limitations are especially serious if there are major changes in the conditions of the market place. Most conceptual frameworks are constructed for a particular set of marketing conditions. If these conditions change lists cannot help us to anticipate these changes or to understand why they occur. In order to have some reasonable ideas about how changed circumstances might cause things to come out differently, we need to understand why things occur. Theoretically informed research produces results which occur for known reasons. It suggests why things come out one way rather than another.

The second, primarily theoretical, approach to the study of consumer choice also has important limitations. For example, the results of economist's theories of consumer choice depend on numerous propositions about the behavior of the various economic actors. These propositions are based on several, usually untested, assumptions. For instance, the preference orderings of consumers which lead to their choice behavior are often assumed to be complete and to be characterized by convexity and nonsatiety. Consumer's preferences

are assumed to have these properties. Whether they actually do is left an open question. Do people's preferences really have these various properties? Until we establish that they do, we cannot claim that the theoretical results apply to anything in the real world.

This orientation tends to produce a one-sided concern with theoretical rigor. As a result, important questions about consumer choice, e.g. what does the consumer choose to buy, and how much of it does he purchase, are often ignored because they are theoretically intractable. It is easier to generate a rigorous discussion about alternative commodity bundles than about two G.E. televisions and a Farberware toaster vs. one goat and a used Ford Pinto.

Clearly a more active interplay between those who do serious theoretical work and those that do empirical investigations is needed. Recently, there has been a considerable increase in interest among theoretically oriented researchers in empirical work and in experimental methods of testing their theories of consumer choice. Concern for systematic theory on the part of applied researchers is also on an upturn. With these considerations in mind, we sought participants who were active in either theoretical research on consumer choice or who had attempted experimental tests of those theories. The papers which follow are a partial result of that effort.

EXPERIMENTAL STUDIES OF CONSUMER DEMAND BEHAVIOR: TOWARDS A TECHNOLOGY OF  
 MAKING THE SLUTSKY-HICKS THEORY TECHNOLOGICALLY  
 APPLICABLE TO INDIVIDUAL BEHAVIOR<sup>1</sup>

Raymond C. Battalio  
 Texas A&M University

John H. Kagel  
 Texas A&M University

The methods and results of experimental studies of individual consumer demand behavior with human subjects in controlled economic environments and with laboratory animals in operant conditioning chambers are reviewed. The review focuses on some of the technical problems encountered in conducting the experiments reported. The implications of these problems and their solutions for the efficient design of economic experiments and empirical consumer demand research in general are discussed.

During the past several years our research in consumer demand behavior has included experimental studies with human subjects in controlled economic environments and with laboratory animals in operant conditioning chambers. The goal of these studies has been to begin to provide an observational interpretation system for the Slutsky-Hicks theory of consumer demand (Hicks, 1946; Uzawa, 1960) that is adequate for the technological application of this theory to the behavior of individual decision making units.<sup>2</sup>

The purpose of this paper is two-fold. First, this paper provides an opportunity to review our experimental procedures and some of the empirical results obtained from this research. The review of the empirical results, while focusing on tests of the consistency of the behavior of individual consumers with the Slutsky-Hicks theory, also includes, to indicate the breadth of problems which can be investigated using these experimental methods, a brief review of the results of investigations of several empirical hypotheses which have been of interest to consumer demand researchers using data from the national economy. Second, our review will focus on some of the technical considerations involved in the actual design and execution of controlled economic experiments. More specifically, we present a detailed discussion of some of the technical problems we encountered in the recording and processing of the data from our experiments and the implications of these problems on the efficient design of economic experiments and on empirical research in general. Initial solutions for dealing with the problems in question are also presented.

Experimental Studies of Economic Behavior: Human Subjects

The experimental studies reported on below were conducted in a controlled economic environment. Controlled economic environments are organized systems in which individuals live for continuous extended time periods and receive tokens or points for work performed which are in turn exchangeable for present or future consumption goods.<sup>3</sup> From our perspective controlled economic environments closely correspond to the economists' concept of a closed economic system where tokens are money, token payments for work performed are wages and the exchange rates in tokens for consumers' goods are the prices of these goods.

The token economy in which we conducted our initial experiments was established in a female ward for chronic psychotics at Central Islip State Hospital in New York.<sup>4</sup> The experimental design used was, in part, selected to allow us to investigate whether any member of a Slutsky-Hicks system of demand functions was consistent with the observations of individual's purchases. That is, suppose we observe  $n$  distinct goods bundles,  $x^i$ , actually purchased at prices  $p^i$ ,  $i = 1, \dots, n$ . It can be shown that a necessary condition for the observed price and quantity data to satisfy a system of Slutsky-Hicks demand functions is that:<sup>5</sup>

$$(1) \quad p^0 x^0 \geq p^1 x^1 \ \& \ p^1 x^1 \geq p^2 x^2 \ \& \ \dots \ \& \ p^{m-1} x^{m-1} \geq p^m x^m \quad \text{then} \quad \sim (p^m x^m \geq p^0 x^0)$$

(where the symbols " $\sim$ " and " $\&$ " denote negation and inclusive conjunction respectively). In cases where condition (1) is not satisfied the observed price and quantity data do not satisfy any system of Slutsky-Hicks demand functions; *i.e.*, the collection of purchases observed is inconsistent with the Slutsky-Hicks theory. By virtue of the logical equivalence of the Slutsky-Hicks theory and the theory of revealed preference, a test of condition (1) is also a test of the consistency of consumer purchases with the axioms of revealed preference theory.

In conducting this experiment, for each occasion the store was open a separate record was made of the items purchased (exchange point record) for each individual. These records were made by trained patient observers who sat next to the store "salesman" and recorded each purchase as it was made on specially prepared tabulating sheets. The observers were supervised by the ward psychologist. Store expenditures were later calculated from these records by multiplying quantities of goods purchased by their respective prices. A second, independent record of total store expenditures was obtained by collecting each token handed in at the store, stamping it "token store" and storing it by week spent (token record). Since each token had the consumer's name on it and was not transferable to other patients, counting these tokens provides an independent measure of the value of total store purchases by individual patients. Comparison of these two independent measures of total store expenditures reveals that there were errors in making and reporting of observations of these purchases. In most cases where the two measures were not the same, the exchange point record totals were less than the token record totals. The difference between the two measures was less than or equal to 5 tokens, or less than or equal to 10% of the larger measure, for about 80% of all individual observations. The sources of these differences were errors made in the counting and classifying of physical objects and their subsequent recording; *e.g.*, misclassification of commodities, misclassification of tokens by ownership, denomination and week spent, and mistakes in counting, including the failure to record store purchases at the time they occurred. Independent checks of non-store purchases, *i.e.*, room rent, breakfast, honor card rental, etc., showed infrequent errors to have been made in recording these expenditures.

Although the sources of the token store errors are not unlike those reported in consumer purchase panel reports (Neter, 1970; Sudman, 1964, Sudman and Bradburn, 1973), there are several important differences between the measures of the observational error in the consumer purchase data reported here and the measures of observational error available for consumer panel reports.<sup>6</sup> First, our measure of error is based on two independent measures of the same concept which, if carried out correctly, would give the same value without adjustments to either figure. Measures of error in consumer panel reports rely on indirect measures to check for observational errors.

Consequently, the two measures will not yield the same value without adjustment of the indirect measure even if both measurement procedures were carried out correctly; *e.g.*, the indirect measurement of the error in consumer panel reports through means of adjusted commodity shipment data. Second, our error assessment procedure provides a measure of error for each subject, for each observation period, whereas no such effort has been reported for consumer panel data. Thus, a measure of the error for each individual observation can be used in assessing our test results. A comparison of our measures of errors with those reported for consumer panel data indicates that our data are at least as accurate as data reported from consumer panel groups. For example, Sudman's (Sudman, 1964) assessment of error bias in panel records of food purchases as assessed against adjusted shipment data shows that for over half of the 55 items studied there is a net undercounting of purchases by panel members of 20% or greater.

In assessing the outcome of our test of condition (1) a measure of the observational error was incorporated into the empirical analysis. While this analysis showed that the behavior of 95 percent of the individuals was in agreement with the test condition within the measured accuracy of the data, for about one-half of the individuals failure to account for observational inaccuracies would have led us to conclude, incorrectly, that behavior was inconsistent with condition (1). The data were then investigated to determine the sensitivity of the empirical tests actually conducted to measurement errors. For the two subjects whose purchase patterns were found to be inconsistent with condition (1), the measurement of the error was zero for both of the inconsistent weeks for one subject, while for the other subject measurement of the error was zero and one half token respectively for the two inconsistent weeks. Further, in both cases an error of 10% or more in total expenditures would have been sufficient to be unable to distinguish between observation errors and actual contradictions of the theory. Of the 15 subjects who were judged inconsistent with condition (1) prior to accounting for inaccuracies in observations, for 10 of them errors of 3 tokens or less in measuring total token expenditures, which is less than 10% of average per capita expenditures during the test period, were responsible for our inability to distinguish between observation errors and contradictions of the theory. Further, for at least 5 of these subjects, an error of 1 token, the smallest unit of currency, would have been sufficiently large to prevent such a determination.

From a pragmatic point of view, the cases for which the observational errors were too large to allow us to distinguish between observational errors and actual contradictions of condition (1) represent a loss of information which is potentially quite important with respect to designing further tests and in suggesting modifications to the theory. In view of the large experimental changes in relative prices and the relatively high quality of the data set used (compared to measures of observational error available for consumer purchase panel reports), it is clear that further direct examination of condition (1) requires an improved observational technology. We have made some progress toward developing such a technology (Winkler, Kagel, Battalio, 1974). This technology, employed in a subsequent experiment, relied on a system of internal consistency checks and multiple, independently recorded, measures of the same economic activity. If there were discrepancies in the regular daily check of the records the relevant subject was required to stay back at the end of the day until his records balanced. Using this system the unreconciled discrepancy remaining in the records for a three week period, during which approximately 135,000 transactions of earning and spending took place with an average daily income level of about 500-600 tokens, was only 23 tokens.



Recent advances in technology allow for further improvements in the collection, recording and processing of data. A technology is now available whereby coded information, attached to each commodity prior to placing the commodity in the purchase area, can be read with an optical scanner at the time of purchase. At the time of purchase other information such as date, time or subject number can also be automatically recorded. In addition to reducing the errors in counting, recording and transcribing the data, this system would substantially shorten the time lag between the collection of the primary data and a complete analysis of the results of the experiment.

The importance of collecting data in a form that is suitable for continuous and relatively detailed analysis cannot be overemphasized in conducting economic experiments in controlled environments. In our experiments, where 3,000 to 4,000 observations per day were recorded and the day-to-day problems of managing the environment are quite time consuming, it is often months after the completion of the experiment before these data have been in a form suitable for computer analysis and the analysis been completed. Our experience has shown that a much greater range of empirical data could have been obtained from our experiments at little additional cost if the information obtained from a continuous and detailed data analysis had been available and used to modify the design and procedures of succeeding phases of a study. The empirical outcomes of our experiments have never been (nearly) completely foreseen. The more quickly these unforeseen outcomes are incorporated into the experimental design and procedures, the more efficient the utilization of the scarce resources.

It is clear from the above discussion that specific decisions about the allocation of research resources toward improving the accuracy of the data are made in conjunction with the choice of the experimental environment to be used and the particular hypotheses to be examined. Since any proposed test of one hypothesis against an alternative requires some minimal level of data accuracy to allow for observationally distinguishing between the hypotheses, many of these same considerations are relevant in using non-experimental data. That is, the choice of hypotheses which can be examined using any data set is constrained by the accuracy of the data set used. Although these constraints are seldom mentioned when using non-experimental data in economics, most of which is collected for purposes other than its use in testing scientific economic hypotheses, the failure to do so makes it extremely difficult to gauge the empirical content of the published results. In view of our direct experience (which indicates that simple counting and classification errors often exceed 5 percent) and the other independent studies in this area (e.g., Sudman, 1964; Morgenstern, 1963), we conjecture that in using data where no information is given about the level of accuracy an estimate of an average absolute error of 5-10 percent for the combined conceptual, observational, truncation and rounding errors is an optimistic estimate. In this context, when such data are used to examine empirical hypotheses which have a derived demand for accuracy of greater than this level it should be the researcher's obligation to provide the evidence required to enable the reader to gauge the empirical content of the results presented.

In addition to providing an initial test of condition (1) and direct empirical evidence on the importance of an explicit recognition of the observational technology and a measure of the observational errors, this experimental study also provided empirical data important to modifying the Slutsky-Hicks theory.

An analysis of the individual consumer data indicated that many of the individuals who changed their composition of consumption when relative prices changed, buying more of the less expensive commodity and less of the more expensive commodity, did not return to the original consumption pattern when prices were returned to the initial baseline value (Battalio, Kagel, Winkler, et al., 1974). In these cases, compared to the initial baseline consumption, the subjects purchased a greater percentage of the commodity which had previously been relatively less expensive. This response was symmetric across commodities during the experiment. These results were quite similar to the results of an earlier, independent study by Ayllon and Azrin (1968a,b) where baseline consumption patterns were displaced by offering subjects small quantities of a free sample between baseline periods. Ayllon and Azrin found that purchases of the sampled commodities were consistently greater following the free sample period as compared to before it. The range of commodities they investigated included soda, trips to a fair, listening to music, and popcorn. The residual effects of increased consumption of the sampled commodity were found both for patients who had consumed the goods in the first (baseline) phase of the experimental period as well as for those who had not. Ayllon and Azrin's results suggest that the changes in consumption associated with the relative price changes in our experiment may account for the residual effects noted above. From these studies it remains indeterminate whether these sampling effects represent temporary, transitional effects following price changes and sampling periods or whether they represent permanent changes in consumption patterns.

While our primary research aim is to characterize the behavior of individual consumers with a specific class of systems of demand functions, thereby removing much of the uncertainty now present with respect to the effect of a price or income change on an individual's purchases, the relatively detailed micro-data set obtained also provides the basis for a relatively inexpensive examination of some empirical hypotheses which are of interest to consumer demand researchers using data from the national economy. For example, using the above data we have examined the following oft-mentioned empirical hypothesis:

If the individuals in an economy can be characterized by a Slutsky-Hicks system of demand equations, then the aggregate percapita composite demand functions are also characterizable by a Slutsky-Hicks system of demand functions, providing the relative prices of each composite commodity remain exactly proportional.

As the data in Figure 1 illustrate, it is possible to organize these data quite well with the system of demand equations

$$x_{ti} = A_{it} M_{ti}^{-1}, \quad i = 1, 2, 3; t = 1, \dots, 7.$$

A formal test of this system of demand equations has shown that, modifying the system to account for the sampling effects of price changes, the discrepancy between the predicted data points and the actual data points is less than the measured errors of observation. Consequently, the introduction of more complicated systems of demand equations at this point, *i.e.*, systems for which  $\partial x_i / \partial p_j \neq 0$ , is unnecessary (Basmann, Battalio, Kagel, 1974).

Since the change in relative prices also resulted in changes in real wages for some of the consumers, this data set could also be used to examine the shape of the individual's labor supply schedule. Performing this analysis showed that the low income earners responded to an increase in real wages primarily by increasing real income and working about the same while the higher income earners responded with small increases in real income and

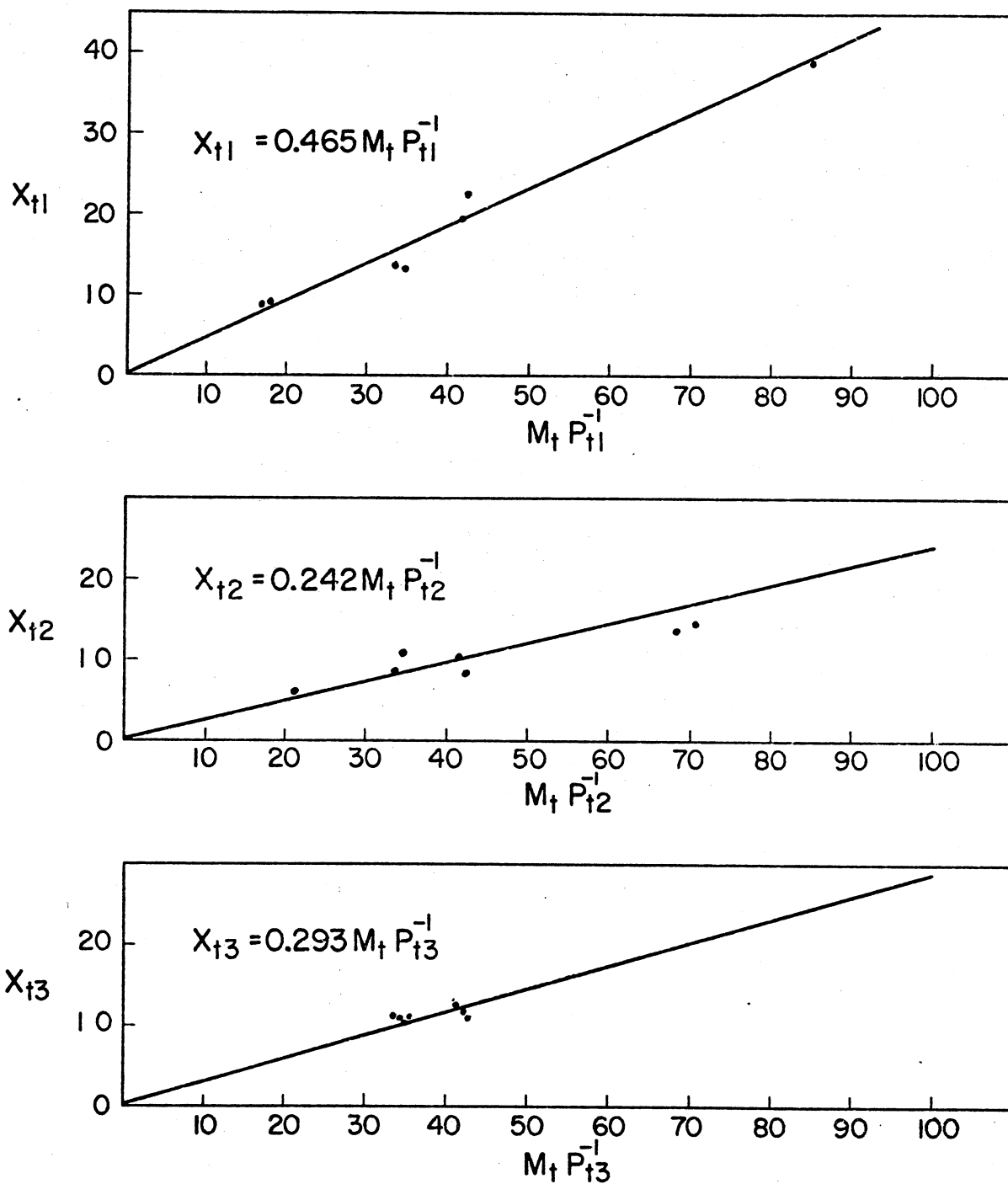


Figure 1. Aggregate per capita income and expenditure data for the Central Islip token economy.

NOTE: The vertical axis designates aggregate per capita purchases of each of the three commodity groups in the economy and the horizontal axis measures aggregate per capita income times the inverse of the price index for each of the commodity groups. Equations for the straight lines determined by the method of least squares.

substantial decreases in work (Kagel, Battalio, Winkler, et.al., 1974). The contrasting behavior of the upper and lower income earners has been accounted for in terms of the initial set of consumer's goods attained relative to available set of consumer's goods. That is, the low income earners had considerably more range, in terms of consuming new goods and increasing consumption of the goods already being consumed, to increase consumption than did the upper income group given the restricted range of consumer's goods available in the economy. (In a subsequent study where the subjects could transfer unspent tokens to currency at the end of a 70 day experiment the high income earners responded to an increase in money wages with little or no decrease in work.) While others have explained labor supply behavior on a similar basis this explanation has heretofore received little direct empirical examination. In this context it is of interest to note that the labor supply response obtained for the low income earners is consistent with the results obtained in the recent income maintenance experiments where the subjects were selected from the working poor (U.S. Dept. H.E.W., 1973).

The preceding discussion has primarily drawn on data from one study in a therapeutic token economy. Controlled economic environments, using volunteer subjects from the national economy who have agreed to reside continuously in these environments for extended periods of time (up to three months), have been developed and successfully applied to the explicit study of particular socio-economic problems (Bigelow and Emurian, 1974; Miles, et al., 1972). While such controlled environments are expensive and require a research team to manage, the relative costs, when weighed against the benefits, make these environments ideally suited to examining certain aspects of economic behavior.

#### Experimental Studies of Economic Behavior: Laboratory Animals

Ever since Darwin it has been widely recognized that behavior as well as structure varies continuously across species and that the applicability of behavioral principles does not stop suddenly at the boundary separating humans from other animals. The principles of economic behavior would be virtually unique among behavioral principles if they did not apply, with some variation of course, to the behavior of non-humans. In comparison to the experimental studies with human subjects reported above, experimental studies of economic behavior using laboratory animals have several advantages. Since the subjects are housed in an experimental chamber and automatic programming equipment is used to carry out and record the transactions, the errors associated with making the transaction and recording the price and quantity observations are greatly reduced at a fraction of the cost of a comparable technology for human subjects. Working with laboratory animals also allows us to conduct experiments and change economic parameters in ways that would be unacceptable on ethical and/or legal grounds with human subjects. For example, although it would be technically feasible to reduce the number of goods in a human consumer's budget set in a token economy, thus simplifying the problems of characterizing a commodity, we do not consider this a practical strategy in view of the empirical evidence that suggests such a simplification may be associated with extensive physiological and psychological changes, many with adverse therapeutic effects. Further, in working with laboratory animals we can, ethically and legally, control for a substantially greater number of environmental factors affecting the organism's behavior than we can with human subjects. This control, e.g., placing an animal in a sound attenuated test chamber, allows a reduction in experimental "noise" to a negligible magnitude compared to studying consumer behavior in controlled economic environments or in national economic systems.

In what follows we summarize the results of two series of experimental studies of consumer demand behavior we have conducted using white male albino rats as subjects (Kagel, Battalio, et. al., 1974). In one experiment the commodities over which the budget set was defined consisted of root beer and Tom Collins mix with food and water available at all times in unlimited amounts. In the other experiment the commodities in the budget set consisted of food and water with no other consumption commodities available outside the budget set. In conducting the experiments our primary goal has been to determine whether laboratory animals would change consumption patterns in response to changes in the budget set and whether the behavior observed could be characterized by members of the class of Slutsky-Hicks demand functions.

In the root beer-Collins mix experiment the subjects were individually housed in a completely enclosed experimental chamber continuously, 7 days a week, except for a period each morning when the cage was cleaned, commodities were replenished and the data were recorded. Two metal levers, projecting into the chamber and located on the back panel of the chamber wall, operated separate dipper feeder mechanisms located behind the panel on which the levers were mounted. A single response on the left lever resulted in the presentation of a dipper cup containing root beer at floor level directly below the lever. A single response on the right lever produced Collins mix from the other dipper feeder. Each subject had a limited sum of lever presses available for operating the dipper mechanisms each day which could be distributed in any way chosen between the two fluids. Under baseline experimental conditions each subject was allotted 300 lever presses with a .05 ml. dipper feeder cup on both dipper feeders giving the budget set represented by the area on or below the straight line A in Figure 2; *i.e.*, the subject could choose any combination of root beer and Collins mix represented by a point on or below line A. Changing the total number of lever presses allotted while maintaining constant dipper cup sizes and lever pressing requirements, would result in shifting out or in of the boundary of the budget set parallel to itself. This gives the same effect that an income change would have on a human consumer's budget set. Thus, the total lever presses allotted to the subject was operationally equivalent to his income level and will be referred to as such. Increasing (decreasing) the size of a dipper feeder cup (and, proportionately, the amount of time the cup remains available) while maintaining a constant lever pressing requirement per dipper presentation, would have the same effect on the rat's budget set as a price decrease (increase) would have on a human consumer's budget set. Thus, the number of lever presses required to obtain 1 ml. of either fluid, maintaining constant lever pressing requirements per dipper presentation, was operationally equivalent to the price of the commodity and will be referred to as such.

Each experimental condition (constant set of prices and income) was maintained for a minimum of 14 days. In addition, a stability criteria of 10 consecutive days, during which no trend was observed in either consumption patterns or time taken to exhaust income, had to be met before conditions were changed.

Mean daily consumption, or purchases, for one of the subjects under baseline experimental conditions (budget line A in Figure 2) is designated by the point  $\bar{A}$  in Figure 2, and the brackets indicate  $\pm$  one standard deviation about the mean.<sup>7</sup> At the income level reported, with prices of both commodities the same, this subject consumed about four times as much root beer as Collins mix. Doubling the price of root beer and halving the price of Collins mix, while simultaneously adjusting income so that the subject could continue to consume the same commodity bundles as before (budget line

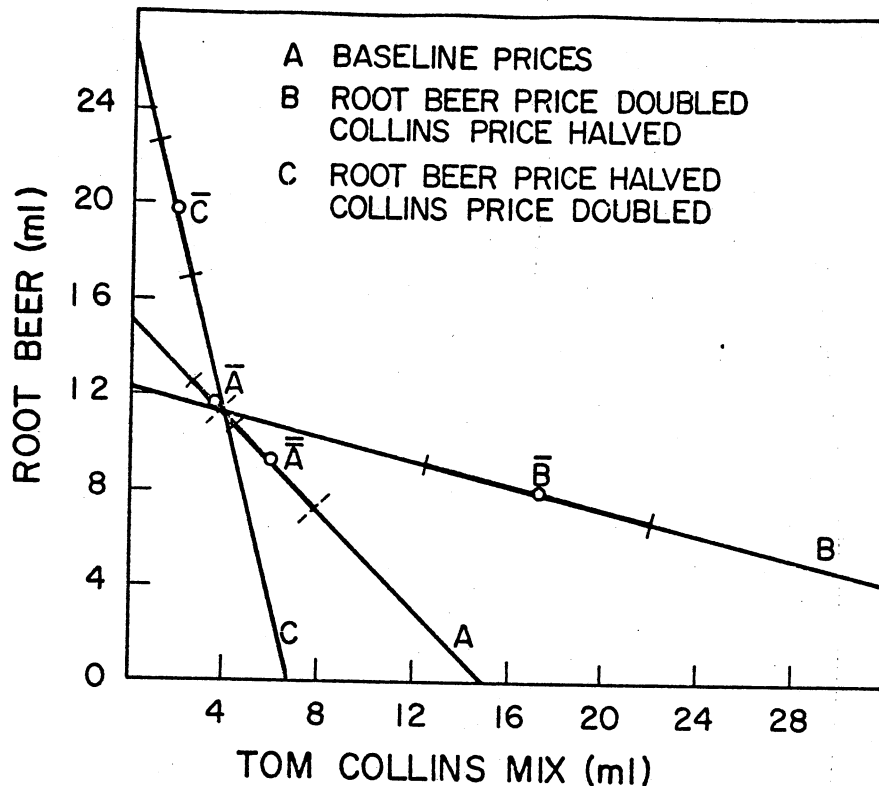


Figure 1.

Budget sets and purchases of root beer and Collins mix under varying prices and incomes. Points on or below each straight line represent the amounts of root beer and Collins mix the subject could afford to purchase under each set of contingencies. Open circles designate mean quantities bought under each experimental condition. Brackets indicate  $\pm$  one standard deviation of the observations about the mean.

B in Figure 2), resulted in the subject consuming about half as much root beer as Collins mix (point  $\bar{B}$  in Figure 2 with the brackets again indicating  $\pm$  one standard deviation about the mean). While the return of prices and income to baseline values resulted in an immediate and sharp reversal in the magnitude of consumption towards the original values (point  $\bar{A}$  in Figure 2), this reversal was incomplete, similar to the sampling effect reported above for human consumers. A further halving of the price of root beer and doubling of the price of Collins mix, while again adjusting income so that the same commodity bundle could be consumed as under baseline conditions (budget line C in Figure 2), resulted in a substantial increase in root beer consumption and reduction of Collins mix intake (point  $\bar{C}$  in Figure 2).

While the above data for mean purchases are consistent with the Slutsky-Hicks theory of consumer demand, the virtually complete control over observational errors maintained in these experiments has allowed us to demonstrate that under a constant set of prices and income the individual consumers do not consume the same commodity bundle from one observation period to another. We view this behavior as resulting from uncontrolled variations in the subject's environment, since even under the most rigorous control of experimental

conditions the subjects are exposed to certain changes in environmental conditions, e.g., changes in temperature, humidity, the composition of commodities, sexual cycles, etc. (Sidman, 1960). Whether the degree of control of experimental conditions maintained in any particular study is acceptable can, of course, only be determined with respect to the hypothesis being examined and the magnitude of the changes in the recorded behavior associated with the changes in the experimental conditions. In the root beer-Collins mix studies the daily variation in purchase patterns was small relative to the differences in the purchase patterns between experimental conditions which enabled us to unambiguously determine that the subjects did change mean consumption patterns in response to changes in the budget set and that this behavior could be characterized by members of the class of Slutsky-Hicks demand functions (Kagel, Battalio, et al., 1974).

At this stage in our research we have eschewed as premature the hypothesis that the daily variations in purchase patterns under a constant set of prices and income are, in any part, attributable to any inherent randomness on the part of the individual consumers. Transparently, the empirical outcomes of our experiments cannot be affected either by adopting or avoiding of such a definition. Notwithstanding this decision, in finding definite members of the class of systems of Slutsky-Hicks demand equations to characterize the data we have adopted the methodological convention of considering purchases under a constant set of prices and income as random variables, i.e., variables derived from a probability distribution. The economy gained, at this stage of research, by considering the effects on consumer purchases of uncontrolled factors in the subjects' environment within the framework of probability concepts rather than explicitly incorporating hypotheses about these factors into the demand equations is more or less unproblematical.

Irrespective of the proffered explanation of the failure of the subjects to consume the same commodity bundle under a constant set of prices and income, this behavior has important implications for the design of further experiments aimed at testing and modifying consumer demand theory. For example, if the design of a proposed test of consumer demand theory calls for "small" changes in prices and/or income then: (1) enough observations must be collected under each experimental condition to distinguish the effects of the systematic changes in the subject's environment from the effects of the unsystematic factors and/or (2) the amount of unexplained variability in purchase patterns under a constant set of prices and income must be sharply reduced. Of course, considerations such as these only become relevant when observational inaccuracies have been reduced to the point where we can distinguish changes in actual behavior.

The experiments in which the budget set was defined over food and water, with no other commodities available for consumption, followed procedures similar to those reported in the root beer-Collins mix studies. In these studies price changes also resulted in changes in the composition of consumption away from the higher priced commodities and in favor of the lower priced goods, although the responses to the price changes were substantially less than those reported in the root beer-Collins mix studies. More importantly, large rotations in the budget set brought about by large increases in the relative price of food (apparent real income held constant) also resulted in severe disruptions of the rat's behavior with the subjects failing to spend all of their income in the allotted 24 hours and steadily losing weight (notwithstanding the fact that as a result of the income constraint they were at only some 80% of normal body weight at the time the suppressed responding

began). Since the experimentally induced changes in the budget set in the root beer-Collins mix experiment and the smaller changes in the budget set in the food-water experiment resulted in changes in consumption patterns quite similar to those reported for individual human consumers (Battalio, Kagel, Winkler, et al., 1973, 1974), the question arises as to whether certain changes in the budget sets of human consumers may also be associated with disruptive behavior. Although we do not have any direct empirical data at present, there is a substantial body of literature in psychology and sociology relating severe disruptions of human behavior, *i.e.*, suicides and mental disorders, to socio-economic factors such as poverty and unemployment.<sup>8</sup> While it is transparent that the disruptions in the rat's behavior (suppressed responding and weight losses) are not the same thing as suicides and mental disorders in humans, this literature does indicate that severe disruptions in human behavior are associated with economic factors. This suggests that in experimental studies of consumer behavior using human consumers the experimenter must guard against changes in the budget set that may result in severe disruptions of the consumers' behavior and be prepared to alter the experimental design accordingly. Further, experimental studies designed to determine the precise factors underlying the disruptions in behavior reported here must be limited to laboratory animals on both ethical and legal grounds. The generality of this behavioral process to human consumers can then be determined by applying the information gained from these studies to the design of permissible experiments with humans.

#### Conclusions

The research described in this paper represents our first steps in the process of achieving our goal of providing an observational interpretation system for the Slutsky-Hicks theory of consumer demand. While we view the actual outcomes of the experiments reported above as important, when appraised against the aim of our research - making the theory technologically applicable - it is clear that they represent merely the beginnings of a task whose completion will require the work of many researchers for a considerable period of time. Our research aim also makes it clear that we do not entertain the concept of "a crucial test" of an economic theory. Rather, we view the theory, and its subsequent modifications, as simply the starting point for designing our experiments, the results of which will be used to modify and augment the observational concepts, the experimental design and the existing theoretical structure. Since every empirical study, whether experimental or not, is unique, the generalizability of a given empirical behavioral relationship can only be determined by continued empirical investigations. In this context, we view the future role of experimental research as an important complement to non-experimental studies. The particular research method, or combination of methods, used for any given problem will, of course, depend on the nature of the question being examined.

#### FOOTNOTES

1. This research was partially supported by NSF Grant #GS32057 "Interpretation Systems for Empirical Economic Theories with Application to Theorems of Factor and Consumer Demand" (Principal Investigators R. L. Basmann, R. C. Battalio and J. H. Kagel).
2. For a more detailed discussion of the concept of an interpreted economic theory and the importance of focusing more research attention on the empirical interpretation system see Basmann, 1974; Brunner, 1969a,b; Morgenstern, 1963; and Sidman, 1960.



3. See Kagel (1972) for a general discussion of token economies and token economy research. The differences between token economy experiments and previous experimental studies of consumer demand behavior are discussed in Battalio, Kagel, Winkler, et al., (1974) and will not be repeated here.
4. A detailed description of this economy, the experiment performed and the results can be found in the papers referenced, and is not reported here. The same is the case for the other experiments referred to in this paper.
5. See for example Theorem 6, p. 144 ff in Uzawa (Uzawa, 1960).
6. Consumer panel reports have also been used to examine condition (1) (Koo, 1963; Koo and Hasenkamp, 1972; Mossin, 1972). For a discussion of the problems in evaluating the results of these studies see Battalio, Kagel, Winkler, et al., 1973.
7. Data are reported for all but the first three days of an experimental condition.
8. Two of the most important studies in this area are the books by Brenner (1973) and Henry and Short (1954). This research is, of course, not without its critics (see the references cited in Kagel, Battalio, et al., 1974).

#### REFERENCES

- Ayllon, T., and Azrin, N. H. The token economy: A motivational system for therapy and rehabilitation. New York: Appleton-Century-Crofts, 1968(a)
- Ayllon, T., and Azrin, N. H. Reinforcer sampling: A technique for increasing the behavior of mental patients. Journal of Applied Behavioral Analysis, 1968, 1, 13-20 (b)
- Basmann, R. L. Modern logic and the suppositious weakness of the empirical foundations of economic science. Schweizerische Zeitschrift für Volkswirtschaft und Statistik, 1975, in press.
- Basmann, R. L., Battalio, R. C., and Kagel, J. H. Aggregate per capita systems of demand functions: A test of a frequently mentioned empirical hypothesis. NSF Project #GS32057, Technical Report #17, Texas A&M University, (mimeographed), 1974.
- Battalio, R. C., Kagel, J. H., Winkler, R. C., Fisher, E. B., Jr., Basmann, R. L. and Krasner, L. A test of consumer demand theory using observations of individual consumer purchases. Western Economic Journal, 1973, 11, 411-28.
- Battalio, R. C., Kagel, J. H., Winkler, R. C., Fisher, E. B., Jr., Basmann, R. L. and Krasner, L. An experimental investigation of consumer behavior in a controlled environment. Journal of Consumer Research, 1974, 1, 52-60.
- Bigelow, G. and Emurian, H. A programmed environment for the experimental analysis of individual and small group behavior. In C. G. Miles (Ed.), Experimentation in controlled environments and its implications for economic behavior and social policy making. Toronto, Canada: Addiction Research Foundation, 1975, in press.

- Brenner, M. A. Mental illness and the economy. Cambridge, Mass: Harvard University Press, 1973.
- Brunner, K. Some reflections on the state of econometric practice. In Essays in honor of Richard Buchner. Zurich: Verlag Schuttlhess & Co., A. G., 1969.
- Brunner, K. Assumptions and the cognitive quality of theories. Synthese, 1969, 20, 501-525.
- Henry, A. F. and Short, J. F. Suicide and homicide. Glencoe, Ill: Free Press, 1954.
- Hicks, J. R. Value and capital. (2nd ed.) Oxford, 1946.
- Kagel, J. H. Token economies and experimental economics. Journal of Political Economy, 1972, 80, 779-785.
- Kagel, J. H. Battalio, R. C., Rachlin, H., Green, L., Basmann, R. L., and Klemm, R. R. Experimental studies of consumer demand behavior using laboratory animals. Economic Inquiry, 1974, in press.
- Kagel, J. H., Battalio, R. C., Winkler, R. C., Fisher, E. B., Jr., Basmann, R. L., and Krasner, L. An experimental analysis of the effect of changes in real wages on work behavior in a token economy. NSF Project #GS32057, Technical Report #3 (mimeographed), Texas A&M University, March 1974.
- Koo, A.Y.C. An empirical test of revealed preference theory. Econometrica, 1963, 31, 646-664.
- Koo, A.Y.C., and Hasenkamp, G. Structure of revealed preference: Some preliminary evidence. Journal of Political Economy 1972, 80, 724-744.
- Miles, C. G., Congreve, G.R.S., Gibbons, R. J., Marsham, J., Devenyi, R., and Hicks, R. C. An experimental pilot study of the effects of daily cannabis smoking on socio-economic behavior. Paper presented at the 30th International Congress on Alcoholism and Drug Dependence, Amsterdam, September, 1972.
- Morgenstern, O. On the accuracy of economic observations. (2nd ed.) Princeton: Princeton University Press, 1963.
- Mossin, A. A mean demand function and individual demand functions confronted with the weak and the strong axioms of revealed preference: An empirical test. Econometrica, 1972, 40, 177-192.
- Neter, J. Measurement errors in reports of consumer expenditures. Journal of Marketing Research, 1970, 7, 11-25.
- Sidman, M. Tactics of scientific research. New York: Basic Books, 1960.
- Sudman, S. On the accuracy of recording consumer panels: I. Journal of Marketing Research, 1964, 1, 14-20.

- Sudman, S. and Bradburn, N. J. Effects of time and memory factors on response in surveys. Journal of the American Statistical Association, 1973, 68, 805-815.
- U.S. Department of Health, Education, and Welfare. Summary report: New Jersey graduate work incentive experiment. Washington, D. C.: Department of Health, Education, and Welfare, 1973.
- Uzawa, H. Preference and rational choice in the theory of consumption. In K. Arrow, S. Karlin and P. Suppes (Eds.), Mathematical methods in the social sciences, 1959. Stanford: Stanford University Press, 1960.
- Winkler, R. C., Kagel, J. H., and Battalio, R. C. Assessment of error in concurrent records of income and expenditure in token economies. NSF Grant GS32057, Technical Report #15, Texas A&M University (mimeographed), 1974.

THEORIES, FACTS AND THE ANALYSIS OF CONSUMER CHOICE

Raymond Dacey  
College of Business Administration  
University of Oklahoma

This paper considers the role of information in a model of inductive learning and decision making. Specifically, conditions are given under which the classical Bayesian framework is maintained in the face of information which is similar to that provided by advertisers under the regulation of the F.T.C. Maintenance of a specific decision making structure is of relevance to those engaged in advertising, regulation, and, more to the point of this session, the controlled testing of theories of consumer behavior.

This paper considers the role of selected public institutions and the mechanisms employed by these institutions. The approach is to characterize "consumers" as "decision-makers" in a formal way, to posit the general problem of "consumer choice" as a generalized decision problem, and to specify the role of inductive logics employed by these decision-makers in interpreting "facts" announced by producers together with "theories" announced by public institutions which regulate the producers' announcements.

Consider the consumer of household products. The F.T.C. specifies and enforces the use of measurement procedures to be applied to the contents of these products and the public disclosure of these applications. Alternatively, consider the investor in publically offered corporate securities. Here the S.E.C. and the Financial Accounting Standards Board are among institutions which specify and enforce measurements of and disclosure about the securities.

In these and other analogous cases, the "governing" institution is imposing a "theory" upon the producers governed. This point is simply a variant of the well-known view that measurements (in general, observations) are "theory-laden."<sup>1</sup> In this paper the disclosed measurements are regarded as "facts" to be interpreted by the decision-maker in light of the imposed theory. Together, these facts and theories bear upon the consumer's choice among alternative products.

It is in this context that the present paper considers the behavior of the consumer. Specifically, the consumer is viewed here as a decision maker in an uncertainty situation. The point of concern is the effect of the receipt of factual and theoretical information upon the consumer's behavior. The major finding is that given a specific and common form for the theory provided by a regulatory institution the consumer, though capable of inductive learning, will, under certain conditions, not revise his view of the world (i.e., his outcome mapping). Further, the conditions vary depending upon the consumer's "inductive philosophy." Put another way, the present analysis provides restrictions on inductive behavior and transmitted information sufficient to guarantee the applicability of classical Bayesian decision theory.

Consider an individual facing a decision problem  $\Delta = \langle X, A \rangle$  with state-space  $X$  and act-space  $A$  who possesses the personality  $P = \langle \rho, \Phi, u \rangle$  where:

- $\rho: X \times A \rightarrow O$  is an outcome mapping,  
 $\Phi: X \rightarrow [0,1]$  is a probability measure on  $X$ ,  
 $u: O \rightarrow \text{Re}$  is a utility function.

Such an individual resolves the problem  $\Delta$  by selecting  $a^* \in A$  so as to maximize  $Eu(a) = \sum u(\rho(x,a))\Phi(x)$ . The point of interest is the individual's reaction to learning something of relevance to  $\rho$  (and perhaps the states of nature).

The connection between outcome mappings and generalizations (i.e., "for all" statements) is as follows. An individual having accepted a generalization possesses the predictive theory whereby the resultant outcome of performing an act in a given state can be determined.<sup>2</sup> For example, given a Keynesian theory of economic behavior (or more accurately, the generalization expressing it), the effect of a tax increase in a state of high unemployment is easily determined. The generalization, i.e., the predictive theory, provides the outcome mapping.

It remains to characterize the individual decision maker's learning process, i.e., his inductive logic. For simplicity the range of the inductive logic is restricted to the decision problem  $\Delta$ . Specifically, it is presumed that the individual has an inductive logic  $I = \langle \underline{P}, P, U \rangle$  where:

- $\underline{P}$  is the class of weak and strong generalizations that yield outcome mappings on  $X \times A$ ,
- $P$  is a probability measure over the sentences of a language  $L$ ,<sup>3</sup>
- $U$  is an expected (epistemic) utility function, defined in terms of  $P$ .

The individual's inductive logic selects, upon the receipt of new information  $\hat{e}$ , a generalization qua outcome mapping so as to maximize  $U$ .

Given a personality  $P = \langle \rho, \Phi, u \rangle$  and an inductive logic  $I = \langle \underline{P}, P, U \rangle$ , both restricted to the decision problem  $\Delta = \langle X, A \rangle$ , the point of interest is the individual's reaction to a sentence  $\hat{e}$  of relevance<sup>4</sup> to  $\rho$ , i.e., a sentence  $\hat{e}$  such that  $P(\rho|\hat{e}) \neq P(\rho)$ . The new sentence  $\hat{e}$  may take either of two forms. First,  $\hat{e}$  may consist of purely factual information to the effect that specific attributive constituents are instantiated by particular individual objects. Second, the new sentence  $\hat{e}$  may be more revolutionary, including factual information and a theory  $T$  (a truthfunctional combination of attributive constituents) introducing and relating a new concept, say  $P_0$ , to the concepts  $P_1, P_2, \dots, P_k$ .

The first case admits of two possibilities. Let  $\hat{\rho}$  denote the outcome mapping selected by  $U$  given  $\hat{e}$ , and let  $\rho$  be the initial outcome mapping. First, if  $\rho = \hat{\rho}$  and  $\Phi(x|\hat{e}) \neq \Phi(x)$  for at least one  $x \in X$ , then the classical Bayesian expected utility framework is applicable and the individual resolves the "posterior" problem  $\Delta$  by selecting  $\hat{a} \in A$  so as to maximize  $E[\bar{u}(a)|\hat{e}] = \sum u(\rho(x,a))\Phi(x|\hat{e})$ . However, if  $\rho \neq \hat{\rho}$ , then the Bayesian framework is inapplicable. The second possibility involves an extension of the Bayesian notion of learning and yields certain interesting peculiarities that are discussed elsewhere.<sup>5</sup>

Let us turn now to the second, more interesting, form of  $\hat{e}$ , i.e.,  $\hat{e} = e \& T$ . The sentence  $e$  provides factual information while the theory  $T$  relates the "new" (monadic) predicate  $P_0$  to  $P_1, P_2, \dots, P_k$ .<sup>6</sup> In order to incorporate the properties of  $e \& T$  into the present analysis it is necessary to specify the properties of  $P$  and  $U$ .

Following Niiniluoto and Tuomela (1973) we take  $P$  to be a Hintikka-type probability measure defined parametrically on a two dimensional continuum with

parameters  $\lambda$  and  $\alpha$ .  $\lambda$ , as in Carnap's one dimensional continuum,<sup>7</sup> is a measure of disorder in the universe;  $\alpha$  is an index of caution.<sup>8</sup> The values of the parameters are specified by the individual, and once specified completely determine his probability measure  $P$ .<sup>9</sup> As such  $P$  is an inductive probability measure and not a purely subjective measure, and thus  $U$ , determined in terms of  $P$ , is an (expected) inductive (i.e., epistemic) utility function. A reasonable form for  $U$  is

$$U(h|e) = P(h|e) - P(h),$$

where  $h$  is any statement in the domain of  $P$  and  $e$  is a sentence of factual information.<sup>10</sup> So defined  $U$  is a measure of relevance and embodies an extension of the concept of substantive information.<sup>11</sup>

For  $\hat{e} = e \& T$ ,  $U$  gives rise to two variants, namely  $U_1(h|e \& T) = P(h|e \& T) - P(h)$ ,

and

$$U_2(h|e \& T) = P(h|e \& T) - P(h|T).$$

The adoption of either  $U_1$  or  $U_2$  to the exclusion of the other reflects the individual's philosophical/methodological position.  $U_1$  seems fitted to the position of a methodological instrumentalist, whereas  $U_2$  seems natural for a scientific realist.<sup>12</sup> The pair  $\langle U, U_1 \rangle$  or  $\langle U, U_2 \rangle$  constitutes the inductive theory of the individual.

Consider now the syntactical structure of  $T$ . The theory  $T$  constitutes a formal (truthfunctional) link between  $P_0$  and  $P_1, P_2, \dots, P_k$ . The simplest such link is a definition. Now definitions, for obvious reasons, should be eliminable and non-creative.<sup>13</sup> However, if a theory  $T$  possesses these properties then it cannot be said to introduce a new concept. Rather,  $T$ , as a definition, merely provides a specification of the "new" concept  $P_0$  in terms of  $P_1, P_2, \dots, P_k$ , such that  $P_0$  serves only as a simplified, economical abbreviation for a (perhaps) complex truth-functional structure built out of  $P_1, P_2, \dots, P_k$ . For example, the concept "calorie" (or "Btu") is an economical substitute for a rather complex conceptual structure involving "temperature," "pressure," and "volume." The question naturally arises: Can a theory  $T$ , if it is an eliminable, non-creative definition of  $P_0$  in terms of  $P_1, P_2, \dots, P_k$  have any effect on the inductive selection of outcome mappings? The somewhat surprising answer is yes.<sup>14</sup>

A theory  $T$  if it is an explicit definition of  $P_0$  in terms of  $P_1, P_2, \dots, P_k$  has a limited inductive role.<sup>15</sup> An explicit definition  $T$  is irrelevant with respect to  $e$  to any generalization  $g$ , i.e.,  $P(g|e \& T) = P(g|e)$  for all generalizations  $g$ . However, if  $T$  is a disjunction of explicit definitions, then  $T$  has inductive import. The notion of  $T$  as a disjunction of explicit definitions is pertinent to the discussion of the governance of measurement and disclosure, as in the instances cited earlier.

Society includes various institutions whose purpose is the pronouncement of definitional theories like  $T$ . For example, the National Bureau of Standards (together with the International Bureau) periodically provides definitional rules concerning the fundamental quantitative concepts of science. In the business world, the Securities and Exchange Commission together with the Financial Accounting Standards Board dictate acceptable accounting procedures for financial reporting by pronouncing definitional theories. Similarly, the Cost Accounting Standards Board dictates acceptable accounting procedures for those engaged in governmental contracting. Of more direct relevance to the

theory of consumer behavior is the Federal Trade Commission, which pronounces definitional theories concerning consumer goods. Very rarely are these theories strictly explicit definitions. Typically, in order to allow some flexibility on the part of those imposed upon, the pronounced theories permit alternative definitions of a concept. For example, neither the concept of "watt" as relates to stereo equipment, nor the notion "octane" for gasoline is explicitly defined. Rather, the pronounced theory  $T$  consists of a disjunction  $T_1 \vee T_2 \vee \dots \vee T_m$ , where each  $T_i$  is an explicit definition (of the new concept  $P_0$  in terms of  $P_1, P_2, \dots, P_k$ ). A user of  $T$  then has the option of employing a specific disjunct  $T_i$ . A consumer may be aware of  $T$  but typically will not know the specific  $T_i$  used to generate the factual information supplied via advertising.<sup>16</sup> Our concern is with the effect of a knowledge of  $T$  upon the consumer/decision maker's selection of an outcome mapping. Specifically, our purpose is to examine the role of  $T$  in the following situation. Suppose, given only the factual sentence  $e$  the decision maker selects outcome mapping  $\rho^*$ . However, given  $e \& T$ , where  $T$  "governs" concepts reported on by  $e$ , he selects  $\hat{\rho}$ . If  $e$  constitutes the decision maker's initial knowledge, then our concern is with the incremental impact of the definitional theory  $T$ . Note that acquisition of  $T$  can be seen as a reduction in ambiguity concerning the concept  $P_0$ . Ambiguity is completely removed if the operative  $T_i$  is made known. Ambiguity can, if it leads to the absorption of  $\hat{\rho}$  in place of  $\rho^*$ , cause the expected utility framework to be inapplicable. Thus, we are also here concerned with the effects of an attempt to reduce ambiguity on the adequacy of a theory of consumer (i.e., decision making) behavior.

Consider now the formal properties of  $T$  as a disjunction of explicit definitions. A concept  $P_0$  is said to be piecewise definable in terms of concepts  $P_1, P_2, \dots, P_k$  in a theory  $T$  if and only if  $T$  logically implies a finite disjunction of explicit definitions of  $P_0$  in terms of  $P_1, P_2, \dots, P_k$ .<sup>17</sup> Clearly, if  $T$  is itself such a disjunction, then it provides a piecewise definition of  $P_0$ .

We now turn to a consideration of the role of the piecewise definitional theory  $T$  in the inductive choice of outcome mappings. More specifically, our interest is in a situation wherein the individual reverses his selection on the basis of  $T$ , i.e., a situation where given the evidence  $e$  and an inductive structure  $\langle U, U_i \rangle$  he selects an outcome mapping  $\rho^*$ , but given  $T \& e$  and  $\langle U, U_i \rangle$  he selects  $\hat{\rho}$  where  $\hat{\rho} \neq \rho^*$ . Thus, our concern is with  $T$ 's role relative to that of  $e$  in reversing the choice of an outcome mapping.

Two results due to Niiniluoto and Tuomela are of immediate interest. Let  $T = T_1 \vee T_2$  (where ' $\vee$ ' denotes 'or') be a piecewise definition of  $P_0$  in terms of  $P_1, P_2, \dots, P_k$ .

- (A) If  $t = T_1 \& T_2$  is incompatible with  $e$ , or if  $t$  is incompatible with the assumption of a non-empty universe, then  $P(g|e \& T) = P(g|e)$ .
- (B) If an individual is cautious and  $e$  reports on a large sample, then  $P(g|e) - P(g) \approx P(g|e \& T) - P(g|T)$ .<sup>18</sup>

Conditions (A) and (B) together with the (necessary) conditions for decision reversal have interesting consequences.

First, consider decision reversal from  $\rho^*$  to  $\hat{\rho}$  for a  $\langle U, U_i \rangle$  individual. Presume that given  $e$  alone he prefers  $\rho^*$  to  $\hat{\rho}$ , i.e.,  $U(\rho^*|e) > U(\hat{\rho}|e)$ . This is equivalent to (RI)  $P(\rho^*|e) - P(\rho^*) > P(\hat{\rho}|e) - P(\hat{\rho})$ . Presume also that

given  $e&T$  he prefers  $\hat{\rho}$  to  $\rho^*$ , i.e.,  $U_1(\hat{\rho}|e&T) > U_1(\rho^*|e&T)$ . But this is (RII)  $P(\hat{\rho}|e&T) - P(\hat{\rho}) > P(\rho^*|e&T) - P(\rho^*)$ .

Now suppose the antecedent of (A) holds. Then  $P(\rho^*|e&T) = P(\rho^*|e)$  and  $P(\hat{\rho}|e&T) = P(\hat{\rho}|e)$ . These equations, together with (RI) and (RII) imply a contradiction. Thus, if  $t$  is incompatible with  $e$ , or if  $t$  is incompatible with the assumption of a non-empty universe, then reversal from  $\rho^*$  given  $e$  to  $\hat{\rho}$  given  $e&T$  is impossible.

Now, consider  $\rho^* - \hat{\rho}$  decision reversal for a  $\langle U, U_2 \rangle$  individual. Given  $e$  alone he prefers  $\rho^*$  to  $\hat{\rho}$ , i.e.,  $U(\rho^*|e) > U(\hat{\rho}|e)$ . This is equivalent to (RI). Given  $e&T$  he prefers  $\hat{\rho}$  to  $\rho^*$ , i.e.,  $U_2(\hat{\rho}|e&T) > U_2(\rho^*|e&T)$ . This is (RIII)  $P(\hat{\rho}|e&T) - P(\hat{\rho}|T) > P(\rho^*|e&T) - P(\rho^*|T)$ . Together with two minor assumptions,<sup>19</sup> the antecedent of (B), (RI) and (RIII) imply a contradiction. Thus, if the individual is cautious and  $e$  reports on a large sample so that  $P(\rho^*|e) - P(\rho^*) \approx P(\rho^*|e&T) - P(\rho^*|T)$  and  $P(\hat{\rho}|e) - P(\hat{\rho}) \approx P(\hat{\rho}|e&T) - P(\hat{\rho}|T)$ , and two minor assumptions are valid, then reversal from  $\rho^*$  to  $\hat{\rho}$  is impossible for a  $\langle U, U_2 \rangle$  individual.

The conditions which block  $\rho^* - \hat{\rho}$  reversal for a  $\langle U, U_1 \rangle$  individual are strikingly different from the analogous conditions for a  $\langle U, U_2 \rangle$  individual. For the former, the mere incompatibility of  $t$  with  $e$  is sufficient. However, this same condition causes no difficulty for a  $\langle U, U_2 \rangle$  individual. On the other hand, if a  $\langle U, U_2 \rangle$  individual employs a factual report  $e$  about the observation of a large number of individual objects and is himself a cautious individual, then his reversal from  $\rho^*$  and  $\hat{\rho}$  is impossible. Furthermore, the conditions sufficient to block a  $\langle U, U_2 \rangle$  individual from a  $\rho^* - \hat{\rho}$  reversal do not have the same power for a  $\langle U, U_1 \rangle$  individual. Thus, the different inductive forms characterized by  $\langle U, U_1 \rangle$  and  $\langle U, U_2 \rangle$  inductive theories are quite distinct, and what constitutes an impossibility for one has little or no effect on the other.

As noted earlier, if  $\rho^* = \hat{\rho}$  and  $\Phi(x) \neq \Phi(x|\hat{e})$  for at least one  $x \in X$ , then the classic Bayesian posterior analysis is applicable to the resolution of the problem  $\Delta$ . Thus, for an instrumentalist, a piecewise definition  $T = T_1 \vee T_2$  is, relative to factual evidence, irrelevant to the selection of an outcome mapping as long as  $t = T_1 \& T_2$  is incompatible with  $e$ . However, for a realist, a piecewise definition  $T$  is irrelevant only if the observation report  $e$  concerns a large number of individual objects and the individual is cautious. If these conditions are met, then classical Bayesian analysis will characterize the individual's behavior.

The foregoing analysis provides conditions sufficient to guarantee that a decision maker will maintain the expected utility model as his decision making framework after absorbing a piecewise definition  $T$ . The conditions are only sufficient and not necessary; furthermore, they are not the only sufficient conditions. Together with various restrictions on probability assignments conditions similar to the antecedents of (A) and (B) yield further conditions sufficient to block  $\rho^* - \hat{\rho}$  reversal. These further conditions are not treated here. Their existence is of course of great relevance to any one conducting experiments on the effect of information on Bayesian decision makers.

Blocking outcome mapping reversal is of special interest to those conducting controlled economic experiments, as in the preceding papers delivered by Kagel and Lloyd. One controls an experiment expressly to avoid non-economic influences on the behavior of the subjects. Information introduced as a catalyst can have such an effect. If, however, the information incorporates



the "correct" kind of piecewise definitions, then the undesired outcome mapping shift can be avoided.

#### FOOTNOTES

1. See, e.g., Hanson (1965), Ch. 1 and (1971), p. 17f, and Simon (1970), p. 17.
2. See Bezdek and Hannon (1974) for an example of an economic theory serving as an outcome mapping in a policy decision problem.
3.  $\Phi$  is P restricted to X.
4. On relevance see Carnap (1962), Ch. 6.
5. See Dacey (1973a), (1973b).
6. There is no loss of generality in considering the introduction of only one new predicate. The ensuing analysis generalizes for any finite number of new predicates.
7. See Carnap (1952).
8. See Hintikka (1966).
9. See Hintikka (1966), and Niiniluoto and Tuomela (1973).
10. See Hintikka and Pietarenen (1966) for the initial presentation of U, and see Niiniluoto and Tuomela (1973), pp. 68-70 for a discussion of U.
11. On relevance see Carnap (1962). For a discussion of substantive information see Bar-Hillel (1964), p. 307.
12. On the instrumentalist and realist properties of  $U_1$  and  $U_2$  see Niiniluoto and Tuomela (1973), p. 70. For an extensive discussion of realism and its relationship to instrumentalism, see Hooker (1974). In short an instrumentalist takes a theory to be merely a device of convenience in organizing data; i.e., a theory is a mere instrument. A realist takes a theory literally, adhering to the view that the theory (more or less accurately) describes reality.
13. On eliminability and noncreativity of definitions see Suppes (1957), p. 153 and Simon (1970), p. 20.
14. The role of theoretical concepts in scientific theorizing has spawned a sizeable literature and involves a number of philosophical issues. See the bibliography in Niiniluoto and Tuomela for references. Also see footnote 15.
15. This is not to say that theoretical terms were introduced via an explicit definition have no import whatsoever. See Hintikka and Tuomela (1970), pp. 44-46.
16. Note that if the user, here the consumer, is aware of the specific  $T_i$  being used by the producer, then  $T_i$ , since it is an explicit definition, is irrelevant to any sentence h, i.e.,  $P(h|e \& T_i) = P(h|e)$ . On the general

issue of the information content of advertising see Nelson (1970) and (1974). The latter announces that "Information is generated by advertising because of consumer power in the product market" (1974), p. 730. This is a weak position. Advertising becomes information when translated or interpreted via a definitional theory. Thus consumer power consists of the power to institute regulatory agencies capable of enforcing adherence by producers to definitional theories.

17. Niiniluoto and Tuomela (1973), p. 57.  
 18. Niiniluoto and Tuomela (1973), pp. 72, 75-76.

19. The two further assumptions are

$$\left| \frac{P(\rho|e) - P(\rho)}{P(\rho|e) - P(\rho)} - \frac{P(\rho|e\&T) - P(\rho|T)}{P(\hat{\rho}|e) - P(\hat{\rho})} \right| <$$

$$\left| \frac{P(\hat{\rho}|e\&T) - P(\hat{\rho})}{P(\hat{\rho}|e\&T) - P(\hat{\rho})} - \frac{P(\hat{\rho}|e) - P(\hat{\rho})}{P(\rho|e\&T) - P(\rho)} \right| <$$

both of which are virtually guaranteed by the meaning of  $\approx$ .

#### REFERENCES

- Bar-Hillel, Y. Language and information. Reading, Mass.: Addison-Wesley Publishing Co., 1964.
- Bezdek, R. and Hannon, B., Energy, manpower, and the highway trust fund. Science, 1974, 185, 669-675.
- Carnap, R. The continuum of inductive methods. Chicago: The University of Chicago Press, 1952.
- Carnap, R. Logical foundations of probability. (2nd ed.) Chicago: The University of Chicago Press, 1962.
- Carnap, R. and Bar-Hillel, Y. An outline of a theory of semantic information. In Bar-Hillel, 1964. Pp. 221-274.
- Dacey, R. Decision making under uncertainty with the information responsive outcome mappings. Proceedings of the 1973 Annual Meeting of AIDS, Boston, 1973a, 208.
- Dacey, R. Inquiring systems and decisions under uncertainty. Proceedings of the 1973 Annual Meeting of AIDS, Boston, 1973b, 60-62.
- Hanson, N. R. Patterns of discovery. Cambridge: Cambridge University Press, 1965.
- Hanson, N. R. Observation and explanation. New York: Harper and Row, 1971.
- Hilpinen, R. Rules of acceptance and inductive logic (Fasc. 22 in Acta Philosophica Fennica), North-Holland Publishing Co., Amsterdam, 1968.
- Hintikka, J. Towards a theory of inductive generalizations. In Y. Bar-Hillel (Ed.), Proceedings of the 1964 International Congress for Logic, Methodology, and Philosophy of Science. Amsterdam: North-Holland Publishing Co., 1965a. Pp. 274-288.

- Hintikka, J. On a combined system of inductive logic. Studia logico-mathematica et philosophica in honorem Rolf Nevanlinna. Fasc. 18, Acta Philosophica Fennica. Amsterdam: North-Holland Publishing Co., 1965b. Pp. 21-30.
- Hintikka, J. A two-dimensional continuum of inductive methods. In Hintikka and P. Suppes (Eds.) Aspects of inductive logic. Amsterdam: North-Holland Publishing Co., 1966. Pp. 113-132.
- Hintikka, J. and Suppes, P. Information and inference. Dordrecht, Holland: D. Reidel Publishing Co., 1970.
- Hintikka, J. and Tuomela, R. Towards a general theory of auxiliary concepts and definability in first-order theories. In Hintikka and Suppes, 1970. Pp. 298-330.
- Hooker, C. A. Systematic realism. Synthese, 1974, 26, 409-497.
- Niiniluoto, I. and Tuomela, R. Theoretical concepts and hypothetico-inductive inference. Dordrecht, Holland: D. Reidel Publishing Co., 1973.
- Nelson, P. Information and consumer behavior. Journal of Political Economy, 1970, 78, 311-329.
- Nelson, P. Advertising as information. Journal of Political Economy, 1974, 82, 729-754.
- Simon, H. A. The axiomatization of physical theories. Philosophy of Science, 1970, 37, 16-26.
- Suppes, P. Introduction to logic. Princeton: Van Nostrand Co., Inc., 1957.

SOME CONCEPTUAL ISSUES IN THE  
TESTING OF THE ECONOMIC THEORY OF CONSUMER DEMAND

Penny Baron and Gerald Eskin  
University of Iowa

Cliff Lloyd  
Simon Fraser University

The economic theory of consumer demand attempts to explain the choice among goods available in the market place based on preferences, prices and the consumer's income. This paper reviews some problems encountered when attempting an empirical test of this theory and discusses some promising methodologies to alleviate some of these difficulties. Some feasible controlled market place experiments are suggested based on these methodological considerations. Specific research sites, treatments, methods of control and measurement techniques are discussed.

How do consumers choose among the goods available in the market place? Although there are many partial answers forthcoming from various theories, only one theory claims great generality and comprehensiveness. It is the economic theory of consumer choice. In this theory consumption occurs in such a way that the selected goods bundle maximizes the consumer's utility subject to prices and income. This theoretical structure has existed more or less in its current form since Slutsky (1915), yet definitive empirical tests have not been made.

Our purpose, here, is to review some problems encountered when attempting to construct an empirical test of the theory and to discuss some promising methodologies to alleviate some of these difficulties. Our approach utilizes the Revealed Preference formulation of the theory and certain treatment manipulations in the context of controlled market place experiments. The treatments discussed combine price changes with income compensation. Under these particular treatment manipulations, practical testing of the theory appears to be feasible.

I

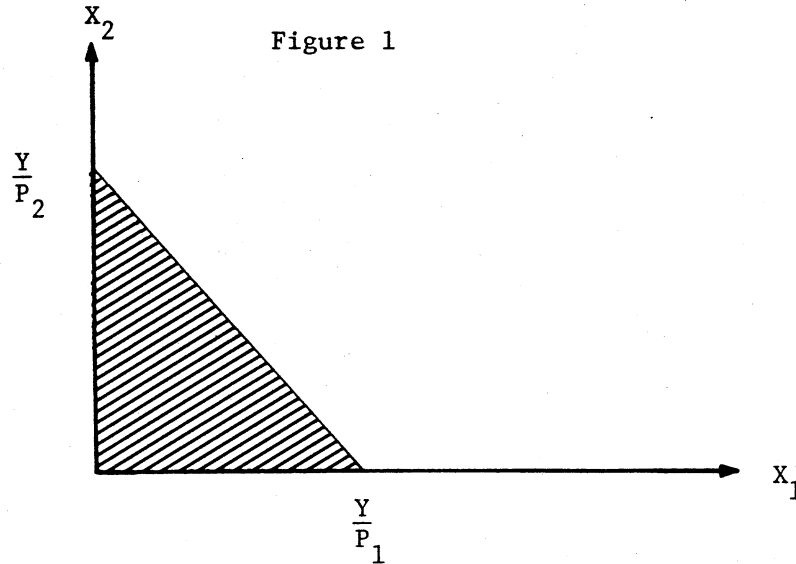
The Traditional Theory of Consumer Behavior

The basic idea of the economist's theory of consumer behavior is that the consumer chooses, from among those things that he can afford, the ones that he prefers. It is supposed that his preferences have a certain regularity of structure (i.e. that he prefers more to less, that if one bundle is preferred to another and the second is preferred to a third then the first is preferred to the third etc.). On the basis of these assumptions a complex and fairly rich characterization of the consumer's behavior may be derived.

Consider for example the case of an individual who is trying to decide how much he will buy of two commodities, the quantities of which we shall designate as  $X_1$  and  $X_2$ . Suppose that the price of commodity 1 is  $P_1$  and that of commodity 2 is  $P_2$ . Then if the consumer had  $Y$  dollars he could just afford any combination of one and two such that the equation

$$P_1 X_1 + P_2 X_2 \leq Y$$

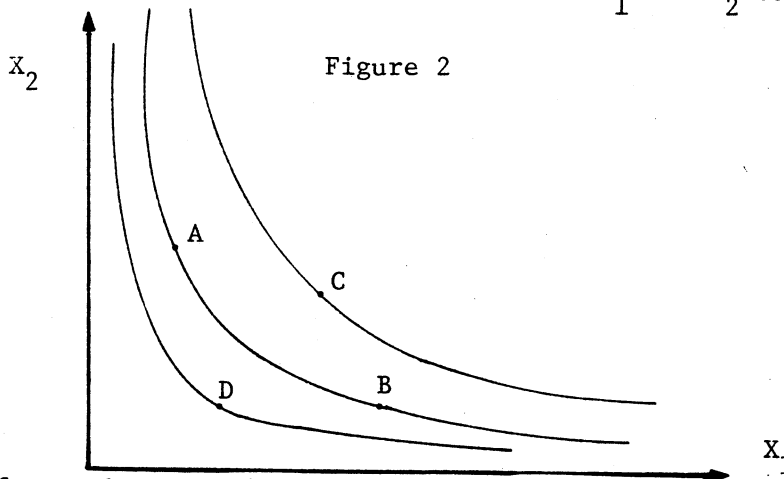
was satisfied. This would be any point in or on the boundary of the triangle in Figure 1.



The triangle is the set of combinations that the consumer can afford and its boundary, the line  $\frac{Y}{P_1}$  to  $\frac{Y}{P_2}$  is his budget restraint representing the equation

$$P_1 X_1 + P_2 X_2 = Y.$$

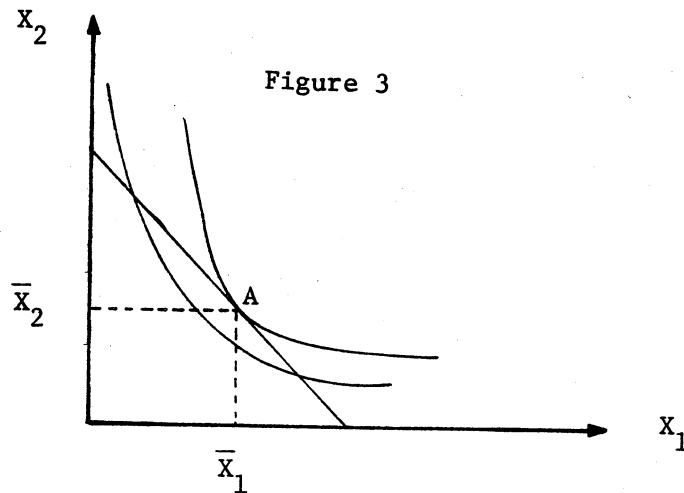
The assumptions that are traditionally made regarding the consumer's preferences imply that these may be represented by a set of indifference curves as in Figure 2. Each such curve represents a set of combinations of  $X_1$  and  $X_2$  such that the consumer



would be indifferent between them. Thus the consumer would be indifferent between points A and B. He would not prefer either of them over the other. Each such curve also divides the entire commodity space (the  $X_1, X_2$  plane) into three regions. One such region is the set of points below the indifference curve. Every point in this region is inferior to every point on the curve. Thus in Figure 2 points A and B are both preferred to point D. Another region is the set of points above the curve, every one of which is preferred to every point on the curve. Accordingly C is preferred to both A and B in Figure 2. The third region is, of course, the curve itself.

The problem of finding the preferred combination of commodities consistent with a given budget is simply identical with that of finding the highest indifference curve that the budget restraint touches. Thus in Figure 3 the consumer will choose

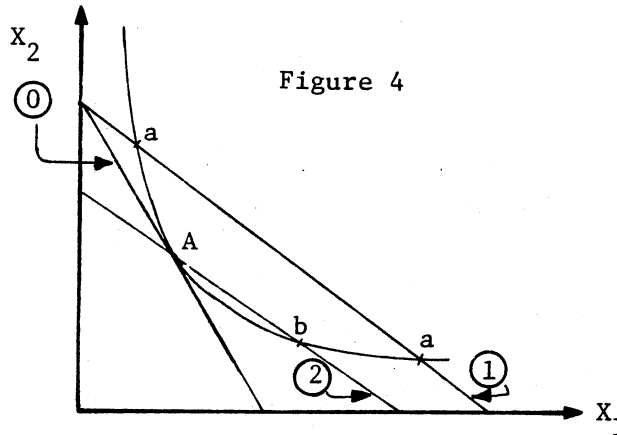
point A consuming  $\bar{X}_1$  of commodity 1 and  $\bar{X}_2$  of commodity 2.



## II

### Empirical Implications of the Theory

To see what sort of implications this theory has, let us consider a consumer confronted with a price change. Suppose that for the price line ① in Figure 4 the consumer chooses the point A. Now suppose that the price of commodity 1 falls so that the budget line becomes less steep like the price line ②. It can easily be seen that the consumer could buy more or less of either commodity after this price change than before it. He will choose some point between the two



points "a" on line ① but which such point our theory does not predict. Now suppose that simultaneous with the fall in the price we had "compensated" the individual for the price change by altering his purchasing power so that he could just buy after the price change what he did buy before it. This would yield the budget line ② in Figure 4. The consumer confronted with this "compensated variation" in price would necessarily choose some point between point A and b on line ② in Figure 4. But every such point is to the right of A. The consumer in this case would have to buy more of the commodity the price of which had fallen or our theory must be wrong.

The choice process described above is presumed to lead to a demand relation

$$(1) \quad X_i = f_i(P_1, \dots, P_i, \dots, P_n, Y).$$

The empirical content of the theory must concern the observable quantities  $(X, P, Y)$  and the function  $f$  which can in principle be determined by the observation set  $(X, P, Y)$ .

It is often assumed that the preference ordering which resulted in the demand function (1) is describable by a utility indicator which is twice differentiable, and convex in all arguments.

It can be shown that

$$(2) \quad \left. \frac{dX_i}{dP_j} \right|_{dY = dP_j X_j} = \frac{\partial f_i}{\partial P_j} + \frac{\partial f_i}{\partial Y} X_j$$

where the expression  $\left. \frac{dX_i}{dP_j} \right|_{dY = dP_j X_j}$  represents the ratio of differentials

$\frac{dX_i}{dP_j}$  for the situation:  $dP_j \neq 0$ ,  $dP_k = 0$  for all  $k \neq j$  and  $dY = dP_j X_j$ . That is

the expression measures the change in consumption of some good  $X_i$  given a price change for good  $j$  under conditions of income compensation of the type illustrated in Figure 4. We shall subsequently write this expression

$$\frac{dX_i}{dP_j},$$

the conditionality  $dY = dP_j X_j$  being implicitly assumed.

It can be shown that certain restrictions on  $f$  and  $\frac{dX_i}{dP_j}$  are implied by the preference (utility maximization) model. For the two good case these restrictions are:

$$(3) \quad \frac{dX_1}{dP_1} < 0, \quad \frac{dX_2}{dP_2} < 0$$

(this is the property being described in Figure 4)

$$(4) \quad \frac{dX_1}{dP_2} = \frac{dX_2}{dP_1}$$

$$(5) \quad P_1 \frac{dX_1}{dP_1} + P_2 \frac{dX_2}{dP_1} = 0$$

$$P_1 \frac{dX_1}{dP_2} + P_2 \frac{dX_2}{dP_2} = 0$$

$$(6) \quad \left( \frac{dX_1}{dP_1} \cdot \frac{dX_2}{dP_2} - \frac{dX_1}{dP_2} \cdot \frac{dX_2}{dP_1} \right) > 0$$

Results (3) through (6) and their generalization to  $n$  variables constitute potentially refutable implications of traditional demand theory. These results have been known, more-or-less since the work of Slutsky in 1915. Still they have

not been conclusively tested. In the section that follows we consider some reasons for this state of affairs.

### III

#### Problems in Testing Consumer Theory

The empirical testing of the theory of consumer demand involves several substantive problems. We discuss below the most important ones for our purpose.

##### Aggregation

The empirical implications of the theory are about the behavior of the individual consumer confronted with price and income changes. The theory has little to say regarding the behavior of groups of consumers. This is not to say that one can not use it to understand the actions of groups, but rather that one can not test it by observations of the behavior of aggregates of people. The reason for this is simply that the income variables relevant to consumption purchases are individual incomes (Hicks, 1946; Lloyd, 1967). It matters who has the money. This means that research based on aggregated data is not directly relevant. [But see Brown and Deaton (1972) for a review of the current state of this literature and its confusions.]

Koo (1963) and Koo and Hasenkamp (1972) have attempted to test the theory using diary records of all food items purchased by individual families in a consumer panel. This procedure avoids much of the aggregation problem but is subject to numerous kinds of reporting errors. Battalio, Kagel et. al. (1973) report a study in a controlled "token economy" which relies on the direct observation of all the purchases by each individual consumer.

##### Required Controls

In order to test the theory, it must be possible to observe the individual's reaction to single price changes for one commodity (or possibly a composite commodity).<sup>1</sup> In a normal environment, without controls, numerous price changes over a variety of commodities will occur on any given day. In such a situation it would be impossible to attribute individual consumption adjustments to particular price changes as the theory requires. It is necessary to be able to control the goods and prices confronting the consumer. Whenever stocks run out, unmeasured price changes occur, or new products are offered for sale, the conditions underlying the predictions of the theory are not met. The consumer's choices in such situations can not be used to indicate whether he is behaving in a way that is consistent with the theory. These considerations suggest that one must either locate an actual economy which can be subject to controls, or build one's own temporary experimental economy. Battalio, Kagel et. al. (1973) chose to do the latter with notable success. The work by Koo (1963) and Koo and Hasenkamp (1972) suffers from a lack of controls over the actual prices and goods confronted by their consumers.

The key consideration with respect to control is that the consumer be presented with a detectable opportunity to behave in a way that is unambiguously inconsistent with the theory. The situation in the usual market economy does not offer such an opportunity. Price changes are not only numerous, but small. Yet Battalio, Kagel, et. al. (1973) have shown that, if attention is given to errors in measuring the consumer's adjustment to these changes, only relatively large price changes produce reliable effects. A single relative price change which is also large, and which occurs in an economic environment where all of the other prices are constant, is not a very likely naturally occurring event. Deliberate experimental manipulation and control are necessary.

##### Requirement of Complete Observations

Some of the implications of the theory (in particular (5) above and the revealed preference notions to be discussed below) require that we observe all of the



individual's purchases. This requirement is particularly difficult to meet if one must rely on records of consumption kept by consumers themselves rather than on direct observations, or if there are numerous places in which a consumer can choose to make his purchases. This consideration suggests that, if we wish to enumerate by consumer, all transactions at the point of sale, then a single purchase site is desirable. This requirement indicates that a very small and isolated economy is required.

### Shifts in Preference

The consumer in a normal environment is frequently confronted with stimuli likely to alter his preferences. Advertisements bombard him constantly. Weather changes alter his clothing, his appetite and his use of appliances. New commodities are introduced into the set available to him. All of these forces are likely to alter his preferences, and results (2) through (6) hold only if preferences are constant. This problem can be mitigated by attempting to isolate the consumer from as many of these confounding influences as possible, and by employing experimental treatments which are robust enough to be little affected by small disturbances.

### Measurement Error

Tests of consumer demand theory are very sensitive to possible measurement errors. Battalio, Kagel, et. al. (1973) found that, when allowance was made for measurement error, many cases which were apparently inconsistent with the theory, could not be distinguished from measurement error. Experimental procedures allow for the control of measurement errors, and the use of direct observations makes it possible to employ multiple measures on each transaction. Only multiple measurements allow for the assessment of magnitudes as well as direction of measurement errors.

### Estimation and Approximation Errors

There are two basic approaches to testing the various implications of the theory (i.e. equations 3-6). The most frequently used method involves estimation of the parameters of some general demand function which has the property that for some values of its parameters equations (3) through (6) are satisfied while for others it is not. In this procedure, estimates are developed, then tested to see if they differ significantly from values that are consistent with the theory.

This approach assumes that a general class of demand functions is known; the problem is simply to discover which of two conditions (consistent vs. inconsistent) prevails. Unfortunately many of the functional forms used do not have sufficient generality. Hence, they impose undesired restrictions on preference and demand. Given these restrictions, when a null hypothesis concerning (3) through (6) is falsified there remains the possibility that it is the improper specification of the model within which estimation is conducted which has caused the result.

If the null is sustained, a specification error could also be responsible. Poorly fitting models tend to produce large error terms. When these estimated errors are used in the statistical tests, bias in favor "no difference" nulls result.

An alternate approach involves the calculation of ratios based on direct observations such as the ratio of demand changes forthcoming from a price change ( $\Delta X/\Delta P$ ) or the demand change forthcoming from an income change ( $\Delta X/\Delta Y$ ). These ratios are considered estimates of the partial derivatives,  $\partial X/\partial P$  and  $\partial X/\partial Y$ . Based on these estimates one can see if conditions (3) through (6) are satisfied.

It is well known that these ratios of discrete changes are not identical to the partial derivatives in most cases. An exception is the linear case,

but unfortunately linear demand functions are not consistent with the utility maximization premise. There is then the possibility that nonsatisfaction of the various conditions in a particular data set could be due to approximation error.

As an example, consider the prospects for testing the negativity of the own substitution effect (equation 3). In order to approximate (2) hence test (3) we would calculate the expression,

$$\left(\frac{\Delta X_1}{\Delta P_1}\right) + \left[X_1 \cdot \frac{\Delta X_1}{\Delta Y}\right].$$

If both  $(\Delta X_1/\Delta P_1)$  and  $(\Delta X_1/\Delta I)$  are negative, our estimate

of (3) will be negative and support the theory. If however opposite signs prevail, the total expression might turn out to be positive. This result might obtain in a situation where there was upward bias in the component with a positive sign or a bias towards zero in the component with the negative sign. An additional ambiguity concerns how to estimate the  $X$  term in this expression. It is not clear whether the estimate should be based on the quantity of good 1 before the price change or the quantity demanded after it. The theory does not tell us because it applies only to an infinitesimally small change in  $P$  and  $I$ , hence to small  $dX$ .

The above problems are inherent in a calculus treatment of a theory in which functional forms are not known. Next we consider a reformulation which does suffer from these difficulties.

#### IV

##### Revealed Preference and Experimental Method as Testing Aids

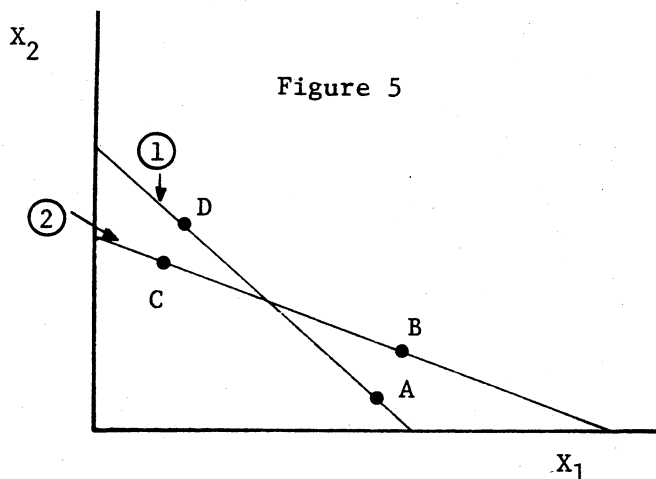
Some of the above mentioned testing difficulties are alleviated by an alternative formulation of the theory, the revealed preference approach. Tests of revealed preference are based on the directly observable quantities  $(X, P, Y)$  as compared to the derivatives of the calculus approach. The usefulness of this "revealed preference" formulation is enhanced by the fact that it is implied by the classical approach. Indeed Uzawa [1960] has shown that satisfaction of the strong axiom of revealed preference is equivalent to the existence of a Slutsky-Hicks system of demand relations.

The basic idea of revealed preference is that if a consumer purchases a goods vector  $X^1$  when he could afford some vector  $X^2$ , then it must be the case that he does not prefer  $X^2$  to  $X^1$ . Also he will not buy  $X^2$  if he could afford  $X^1$ . One might even be willing to assert that not only is it the case that  $X^2$  is not preferred, but that  $X^1$  is preferred to  $X^2$ . This latter formulation is the basis of the strong-axiom of revealed preference. Assuming that  $X^1$  is bought at prices  $P^1$  and  $X^2$  at prices  $P^2$ , a requirement of the axiom is:

$$(8) \quad (P^1 X^1 > P^1 X^2) \Rightarrow (P^2 X^1 > P^2 X^2).$$

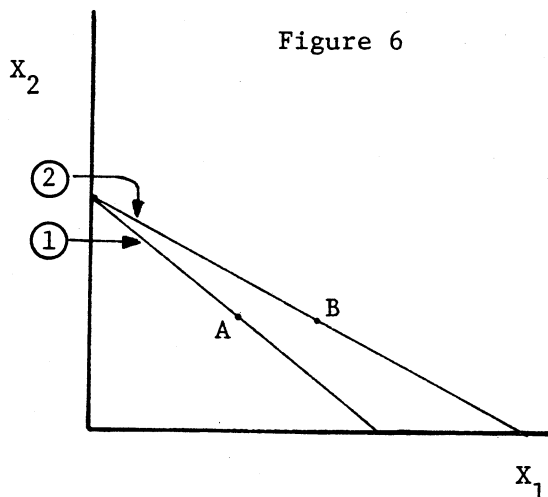
That is, if when the consumer buys  $X^1$  he could more than afford  $X^2$ , then, when he buys  $X^2$  he could not afford  $X^1$ .

Direct test of the theory is possible based on the revealed preference approach. For example, consider a consumer facing the budget line ① in Figure 5 and who purchases the quantities designated by point A. Subsequently, if prices and income should shift in such a way that the new budget line is the one labeled ② then, if he should select the point C, it could be concluded that his actions were inconsistent with the strong axiom of revealed preference. When faced with the original budget ①, point C was available, hence A was revealed preferred to C. Subsequent selection of C violates this preference ordering, hence



falsifying the theory. If B had been selected, falsification could not have resulted. It would merely have established that B is a preferred point. Similarly, original selection of D and subsequent selection of C is consistent (because the consumer could not afford D when C was bought).

The above establishes that if a consumer was faced with income and prices resulting in budget sets (1) and (2) and if we could observe the consumption pattern under these two situations, then there exists the possibility of observing behavior inconsistent with the theory. Thus, conceptually, testing is possible. But just any arbitrary price or income variation will not necessarily provide data for testing. For example, consider a consumer faced with the budget restriction (1) in Figure 6. Now let there be a decrease in the price of  $X_1$  such that budget line (2)



becomes the relevant restriction. Point B is revealed preferred to A by this manipulation but there does not exist a point on budget line (2) such that selection of that point is inconsistent with the revealed preference axiom. Price variations as in Figure 6 help establish the preference ordering but do not, alone, provide tests of the theory.

This suggests, that even abstracting from the numerous difficulties discussed above, actual market data may not always vary in ways that would allow proper testing. In general, the realism of the marketplace will involve relatively small variations in price of a capricious sort. However, when controlled manipulation of prices and income is possible, variables can be manipulated in ways that maximize the strength of the tests. Assuming this possibility exists, what testable propositions can be formulated using the revealed preference approach? The following are suggested:

1. Reversibilities--Consider some change in the budget constraint and corresponding shift from some point A to some point B. Next consider a reversion to the original

budget line. Reversibility requires that the quantity consumed return to point A. This implication follows from the static nature of the theory. Although shifts in preference or lengthy adjustment processes could account for non-reversible behavior, the admission of such arguments renders the theory vacuous (i.e. nontestable). In the research suggested below we seek conditions under which relative stability in preference should hold and argue that if nonreversibility is observed under these conditions, the theory has questionable utility as a guide to empirical analysis using time series data.

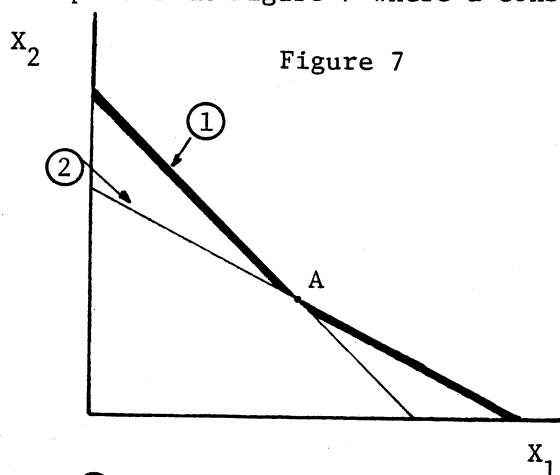
2. Homogeneity--If all prices and income are changed by the same proportion then no shift in consumption should result. This follows immediately from the observation that such manipulations do not alter the budget restriction.

3. Negativity of the Compensated Price Effect--Consider a price change for some good  $X_i$  from  $P_i^0$  to  $P_i^1$ , all other prices unchanged. If an income compensation of the amount  $(P_i^1 - P_i^0) X_i^0$  is given then the strong axiom of revealed preference implies that consumption of good  $X_i$  will fall.

That is

$$(X_i^1 - X_i^0) < 0.$$

The situation is as depicted in Figure 7 where a consumer originally at



point A on budget line (1) must shift to the shaded area of budget line (2) when faced with the new budget situation, a shift to the unshaded area being inconsistent with the theory. Any such shift entails increased consumption of commodity 1. Any consumer, originally at point A on budget line (2) when subsequently faced with budget line (1) must shift to the shaded portion of that line, the unshaded area being inconsistent with the theory. These sorts of compensated price variations then have the following properties. For a compensated price decrease, each consumer must increase his consumption of the good in question. If he does not change the quantity consumed or if he decreases it, inconsistency with the theory is established. With such compensated price variations we are not faced with situations as in Figure 6 where all behavior patterns are consistent with the theory under some pairs of price and income configurations.

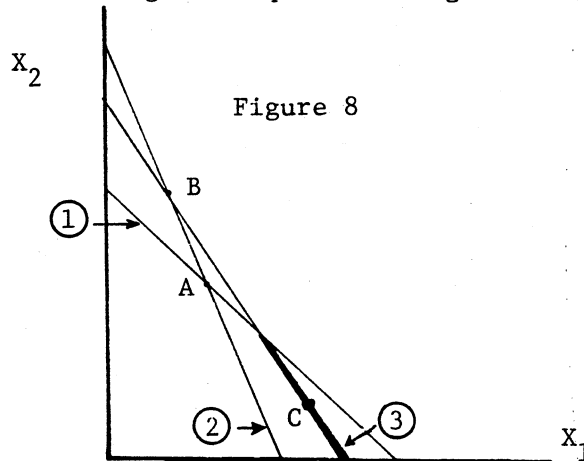
In a sense the compensated price variation maximizes the opportunities for the consumer to behave in an inconsistent manner if the consumer chooses to do so.

The exception to the above argument is the consumer who under the initial set of conditions purchases a zero quantity of the good in question. Here the situation is as in Figure 6. The value of income compensation is zero and inconsistency is not possible.

4. Acyclic Behavior-- The strong axiom of revealed preference requires that if C is revealed preferred to B and B is revealed preferred to A then it can not be the case that A is revealed preferred to C. Behavior satisfying this condition is said to be acyclic.

It is possible to construct a test using income compensation such that, under subsequent pairs of treatment, the negativity of (3.) could hold but acyclicity would be violated.

An example is given in Figure 8. Consider a consumer who purchases a positive quantity of both goods  $X_1$  and  $X_2$ . Under a price increase for good  $X_1$  with appropriate income compensation,<sup>2</sup> if this consumer satisfies (3.), observed consumption will be at a point such as B which will thus be revealed to be preferred to A. Next consider an increase in the price of  $X_2$ , again income compensated, but now from the point B. If the price changes are such that the final budget line (3) intersects budget line (1) in the positive orthant, then three results are possible. Either condition (3.) is violated by the selection of a point to the left of point B or the selected point is revealed preferred to B. If the selected point is revealed preferred to B then it may be the case that it is a point such as C. That is, it is located in the shaded portion of budget line (3). In this case C is revealed preferred to A by the transitive nature of revealed preference. But it is also the case that A is preferred to C because C is in the interior of budget set (1). Experimental manipulations designed to produce budget sets as in Figure 7 and 8



could be employed in experimental tests of the theory. Two operational versions of these tests are as follows:

All commodities available to consumers at retail in an isolated local market are divided into 3 groups ( $C_1, C_2, C_3$ ). Compensated price variations are instituted in the manner described in Exhibit I.

Design I attempts to produce choice situations for consumers like those described in Figures 5 and 7. This design allows for two tests of the negativity of compensated price effect, comparison of purchase behavior under Treatment A with the behavior observed under treatment B and a similar comparison between treatments A and C. (See 3. above.) It allows for one test of reversibility, comparison of purchase behavior under treatment A during Period 1 with the behavior under treatment A during Period 3.

Design II is intended to produce situations like the one described by Figure 8. This design allows for two tests of the negativity effect (3.), AB and BC, and one test of reversibility, comparison of Periods 1 and 4 under treatment A.

Exhibit I  
Price Levels by Commodity Groupings for Each  
Treatment Period  
(Base = Initial Price Level)

Design I  
Commodity Grouping

Treatment in each 4 Week Period	<u>1</u>	<u>2</u>	<u>3</u>
1-A	Base	Base	Base
2-B	Base	+25%	Base
3-A	Base	Base	Base
4-C	Base	Base	+25%

Design II  
Commodity Grouping

Treatment in each 4 Week Period	<u>1</u>	<u>2</u>	<u>3</u>
1-A	Base	Base	Base
2-B	Base	+25%	Base
3-C	Base	Base	+15%
4-A	Base	Base	Base

All price changes are accompanied by appropriate income compensations.

---

Additionally, this design allows one test for acyclic behavior, comparison of purchase patterns under treatments A, B, and C.

Design II seems somewhat more attractive on account of this additional testing possibility. However, its success depends on being able to actually produce a choice situation which looks like Figure 8 for many consumers. Since it requires very rapid adjustment to successive price changes, it is somewhat risky in the absence of information about adjustment lags. Design I is our preferred design for an initial investigation.

V

Testing Procedures

The above discussion has established the general requirements for a test of the theory. Proper manipulations of price coupled with appropriate income compensations have been shown to constitute appropriate test conditions for the theory within the context of a revealed preference approach. Designs for testing the propositions of the theory have been suggested (Section IV). It remains to be shown that the designs proposed are executable and that there exists a site which allows adequate experimental control and at the same time retains some external validity.

It is necessary to discover a "real world" site which allows for the necessary controls, where conditions are stable enough so that consumer preferences do not change, where the economy is closed, where it is possible to observe all market transactions for each consumer in the research population, and where all the relevant conditions of such transactions are controllable or measurable.

Procedures must exist to execute the required design. Adequate measures of dependent variables must exist which yield known and small measurement errors. It must also be possible to execute the required price treatments and income compensation over appropriate time horizons.

Below we briefly discuss some proposed research, currently being considered for funding, which attempts to meet these requirements.

Possible sites include the villages of Davis Inlet, Hopedale, Makkovik, and Postville. All are isolated villages along the northeast coast of Labrador, Newfoundland. Each village is served by a store run by the Newfoundland Government Department of Recreation and Rehabilitation. In each town no significant competition exists for the government store. Each of these stores receives a large shipment of goods by boat in early fall. Shortly thereafter very cold weather sets in and no significant shipments of goods can be sent in until the following summer. Prices are set on a cost plus basis and are constant through the winter.

Postville appears to be the preferred site for initial investigations. It has only the government store, with no competition of any sort. The population is estimated at 140. It is a permanent fishing village with a settled population of Scottish origin. The population earns its income by fishing during the period of open water and receives unemployment compensation during the long winter.

The community is extremely isolated and most villagers do not travel outside it during the winter. Accordingly, by controlling prices in the government store, all prices faced by the local consumers will be controlled. Permission has been obtained from a representative of the Newfoundland Government to use the government store as a test site.

Testing will be conducted during the period January 1 through April 31. This will avoid Christmas buying and allow completion of the study well ahead of the spring thaws. During this period the weather in the area is consistently very cold. Accordingly no taste changes due to seasonal variation should occur. There is no advertising done in the local market and the village is sufficiently isolated to be beyond the reach of external advertising. No shifts in consumer preferences will thus arise from advertising efforts. Obviously no new products will be introduced into the environment during its period of more-or-less complete isolation. Procedures are available for dealing with anticipated minor sources of external effect.

Treatments. The variable to be manipulated in this experiment is the price of certain goods with income compensation. Other relevant variables can either be controlled or measured.

Approximately 60 goods which are frequently purchased and which account for a reasonably large proportion of the consumers bi-monthly incomes will be selected for experimental manipulation. The selected goods will be divided in half in such a way that virtually everyone makes purchases within each grouping and that expenditures are approximately equal. Price manipulations will be made over each of these two product groupings. All the remaining goods will constitute a third grouping whose prices will remain constant throughout the course of the experiment.

Timing. The treatment periods will be one month long. This will allow completion of the experiment in the time available. Each treatment period covers two bi-monthly pay periods which should allow some measurement of time effects.

Income Compensation. A continuous record of each consumer's purchases will be maintained via cash register tapes and records from tear off price tags. At the time of a price change each consumer's expenditure, during the previous period, for items whose prices are to be raised, will be determined exactly and the

consumer will be given 15% or 25% of this amount depending on the treatment requirements. Consumers will be told at the start of this study that the general interest of the study is to gain a better understanding of how retail transactions are carried out and that some income adjustments to maintain consistent buying power may occur should prices go up.

Dependent Variable. The dependent variable is the purchase behavior of individual consumers under the various treatment conditions discussed above.

Measurement Error. Direct observations of purchases will be used. Register receipts and inventory tags are the primary records. Errors should be minimal and will be calculated separately for each consumer. The size of measurement error can be determined by comparison of the two separate methods of measuring transactions. Since records are obtained directly during the transaction of a sale, many types of bias have no opportunity to arise. The records themselves should also suffer from minimal error since, in one case, the record is the actual cash register receipt which displays a completely itemized listing of each purchase, and, in the other case, it is the coded tags from each item purchased.

Data Analysis. The analysis procedure is straightforward. The expenditure pattern of each consumer will be examined for changes in relative amounts spent for each commodity grouping across treatment periods. Compatibility of these expenditure patterns with revealed preference theory is verifiable using the obtained data on prices and quantity purchased. The number of consumers who reveal themselves to be consistent with respect to each proposition being tested will be determined.

#### FOOTNOTES

1. A composite commodity is a group of commodities the prices of which always vary by the same percentage.

#### REFERENCES

- Battalio, Raymond C., Kagel, John H., Winkler, Robin G., Fisher, Edwin B. Jr., Bassman, Robert L., Krasner, Leonard. A test of consumer demand theory using observations of individual consumer purchases. Western Economic Journal, 1973.
- Brown, J.A.C. and Deaton, A.S. Surveys in applied economics: Models of consumer behavior. Economic Journal, 1972, 82, 1145-1236.
- Hicks, J. R. Value and capital. (2nd ed.) Oxford: Clarendon Press, 1946.
- Kagel, J. H. Token economics and experimental economics. Journal of Political Economy, 1972, 80, 779-785.
- Koo, Anthony Y. C. An empirical test of revealed preference theory. Econometrica, 1963, 31, 636-664.
- Koo, Anthony Y. C., and Hasenkamp, George. Structure of revealed preference: Some preliminary evidence. Journal Political Economy, 1972, 80, 724-744.
- Lloyd, Cliff. Microeconomic analysis. Homewood, Illinois: Richard D. Irwin, 1967.



- Marimont. A new method of checking the consistency of preference matrices. Journal of the Association for Computing Machinery, 1959, 6, 67-71.
- Neter, John. Measurement errors in reports of consumer expenditures. Journal Marketing Research, 1970, 7, 11-25.
- Slutsky, E. E. Sculla teoria del bilancio del consumatore. Giornale degli Economisti e Rivista di Statistica, 1915, 51, 1-26.
- Uzawa, H. Preference and rational choice in the theory of consumption. In K. Arrow, S. Karlin and P. Suppes (Eds.), Mathematical Methods in the Social Sciences. Stanford: Stanford Press, 1960.

AN EXPERIMENTAL STUDY OF THE RELATIONSHIP BETWEEN  
RESPONSES TO PRICE CHANGES AND THE PRICE LEVEL FOR SHOES

Nonyelu G. Nwokoye<sup>1</sup>  
University of Massachusetts

The responses of subjects to shoe price changes were studied using the concepts of price tolerance and discount limit. Price tolerance was defined as the largest price increase that would make a shopper still consider purchasing; discount limit was defined as the largest price decrease that would make the shopper still believe the original price. While price tolerance generally increased with price level, the increase was nonlinear but appeared to be highly sensitive to the ending digit of the price. There was strong evidence that discount limit increases linearly with price. Price tolerance was significantly less than discount limit for most price levels, indicating that subjects were more sensitive to price increases than to price decreases.

How do consumers react to price changes? Are consumers more sensitive to price increases than to price decreases? How much of a price change is necessary to invoke a behavioral change? Reliable answers to the above questions have practical implications for price-setting.

If intuition suggests that a \$1 price increase is more apparent on a \$10 item than on a \$100 item, does this imply a systematic relationship between the perception of price increase and the price level? Weber's law has often been invoked when researchers discuss perception of price changes. Weber's law states that the increment in stimulus intensity needed to produce a just noticeable difference (JND) in response is directly proportional to the stimulus.

Researchers disagree on whether Weber's law could be assumed in pricing and response relationships derived from it (see Kamen and Toman, 1971; Monroe, 1971; Gabor, Granger, and Sowter, 1971). The major problem is that Weber's law cannot be directly tested for price, because, as has been articulated by Stapel (1972), everybody can notice even a one-cent difference. This suggests that there is no JND for price, and that experimental methods which depend upon the JND concept would not be directly applicable in pricing studies. One way out of the problem is to operationally define the smallest change in price that will induce a well-defined behavioral change in the buyer.

This paper reports the results of a study in which changes in price for shoes (plus or minus about a level) obtained from subjects' responses were related to price level over a wide range of prices. Directional effects were also examined to see if price sensitivity is different for price increases as compared to price decreases. Such differences have been reported by other workers using other methods (e.g. Pessemier, 1960; Uhl, 1970).

## Hypotheses

Price increases and decreases are first defined and given particular labels. For price increases from a level, price tolerance is defined as the largest increase in the price of a particular product offer that will make a shopper still consider purchasing that product offer. For price decreases from a level, discount limit is defined as the largest decrease in the price of a particular product offer that will make a shopper still believe the initial price. The above definitions imply that the shopper knows and accepts the initial going price of the product offer.

The following hypotheses are set up:

Hypothesis 1: For a given market segment and product class, price tolerance increases linearly as price level increases.

Hypothesis 2: For a given market segment and product class, discount limit increases linearly as price level increases.

Hypothesis 3: For a given market segment and product class, price tolerance is less than discount limit for each price level.

Hypotheses 1 and 2 are suggested by analogy to Weber's law. Hypothesis 3 derives from the fact that price is a sacrifice a buyer makes for acquiring a product, and it is assumed that a buyer would want a price increase (if one must be made) to be as small as possible, and a discount (if one is promised) to be as large as possible.

## Method

Using the product class shoes, mean price tolerance and discount limit were determined for 180 University of Massachusetts undergraduate students in a laboratory setting. Two completely randomized designs were used -- one design for price increases and the other for price decreases. In each design, price treatments were 15 price levels ranging from \$3 to \$50 for a pair of shoes, and 6 subjects responded for each price level. Since the product class was being studied, no particular type of shoe was specified. Each subject was presented with a scenario which, for both price increases and decreases, began as follows:

Shoes come in a wide range of prices, depending on the type and quality of shoe.

Suppose you are out one day shopping, and as you pass by a shoe store you recall that for a while you have been thinking of buying a new pair of shoes. Then you enter the store and find a pair of a certain type and quality that you want. Assume that the price of the shoes is \$P, and it is a price you can pay. However, on this day you do not plan to spend the \$P on shoes, so you leave the store and decide to return a few days later to pick them up.

For price increases, the scenario ended thus:

When you come back to the store you find that the price of the shoes has been increased. What is the highest price you will still consider paying? Assume that your mind is not rigidly fixed on having that particular pair.

For price decreases, the scenario ended thus:

When you come back to the store you find that the price of the shoes has been reduced. What is the minimum price that will make you still believe the original price?

Each subject was given an individual price scale on which a mark was to be made. For price increases, the scale began from the initial price and increased in steps of \$0.25 to a terminal price at least 200 percent higher for initial prices in the range \$3 to \$15, and to at least 50 percent for initial prices in the range \$18 to \$50. For price decreases, the scale began from the initial price, decreasing in steps of \$0.25 to \$0 for initial prices in the range \$3 to \$15, and to a terminal price at least 50 percent lower for initial prices in the range \$18 to \$50. In all cases, the price scale was wide enough for any respondent.

The scenario and price scale were contained on one sheet of paper. The 180 sheets (6 increases and 6 decreases about 15 price levels) were randomly stacked using a random number table and then serially handed out to the subjects.

Since prices were unambiguously specified, problems of sensitization could be serious if repeated measures of price tolerance and expected discount were attempted on the same subject. Thus, each subject responded to only one price change.

#### Data Analysis and Results

Price tolerance or discount limit was defined as the absolute difference between the initial price and the subjects' price response, i.e.

$$|P_i - P_r|$$

where  $P_i$  is the initial price

$P_r$  is the price response

Price tolerance was obtained for each subject responding to a price increase, and discount limit for each subject responding to a price decrease. For each price level, mean price tolerance was computed for the group of subjects responding to increases from that price and mean discount limit was computed for the group responding to decrease from that price. These are shown in Table 1. Mean price tolerance as a percentage of price and mean discount limit as a percentage of price were computed for each price level and are also shown in Table 1. Acceptable price increases ranged from 5 percent for a price level of \$45 to 64 percent for a price level of \$3, with most values below 15 percent. Believable price decreases ranged from 17 percent to 47 percent.

One-way ANOVA was performed on both sets of data. F-tests showed that price level was not significant in explaining the variation of price tolerance ( $p > 0.10$ ) but price level was significant in explaining variation of discount limit ( $p < 0.005$ ).

Figure 1 shows price tolerance as a function of price. There is a wide scatter of points, making the hypothesized linear relationship unlikely to hold. However a simple linear regression fit produced the equation:

TABLE 1

Mean Price Tolerance and Mean Discount Limit and  
Their Percentages for Fifteen Price Levels

Price Level	Mean Price Tolerance (N = 6)	Percentage Price Tolerance	Mean Discount Limit (N = 6)	Percentage Discount Limit
\$ 3	\$ 1.917	63.9 %	\$ 1.417	47.2 %
5	1.000	20.0	2.083	41.7
8	0.625	7.8	2.375	29.7
10	1.000	10.0	3.250	32.5
13	1.583	12.2	4.833	37.2
15	2.208	14.7	4.833	32.2
18	1.833	10.2	3.125	17.4
20	3.917	19.6	4.583	22.9
23	2.833	12.3	5.000	21.7
25	2.083	8.3	6.792	27.2
30	3.792	12.6	6.667	22.2
35	2.417	6.9	8.083	23.1
40	3.083	7.7	12.125	30.3
45	2.250	5.0	8.500	18.9
50	3.875	7.7	12.833	25.7

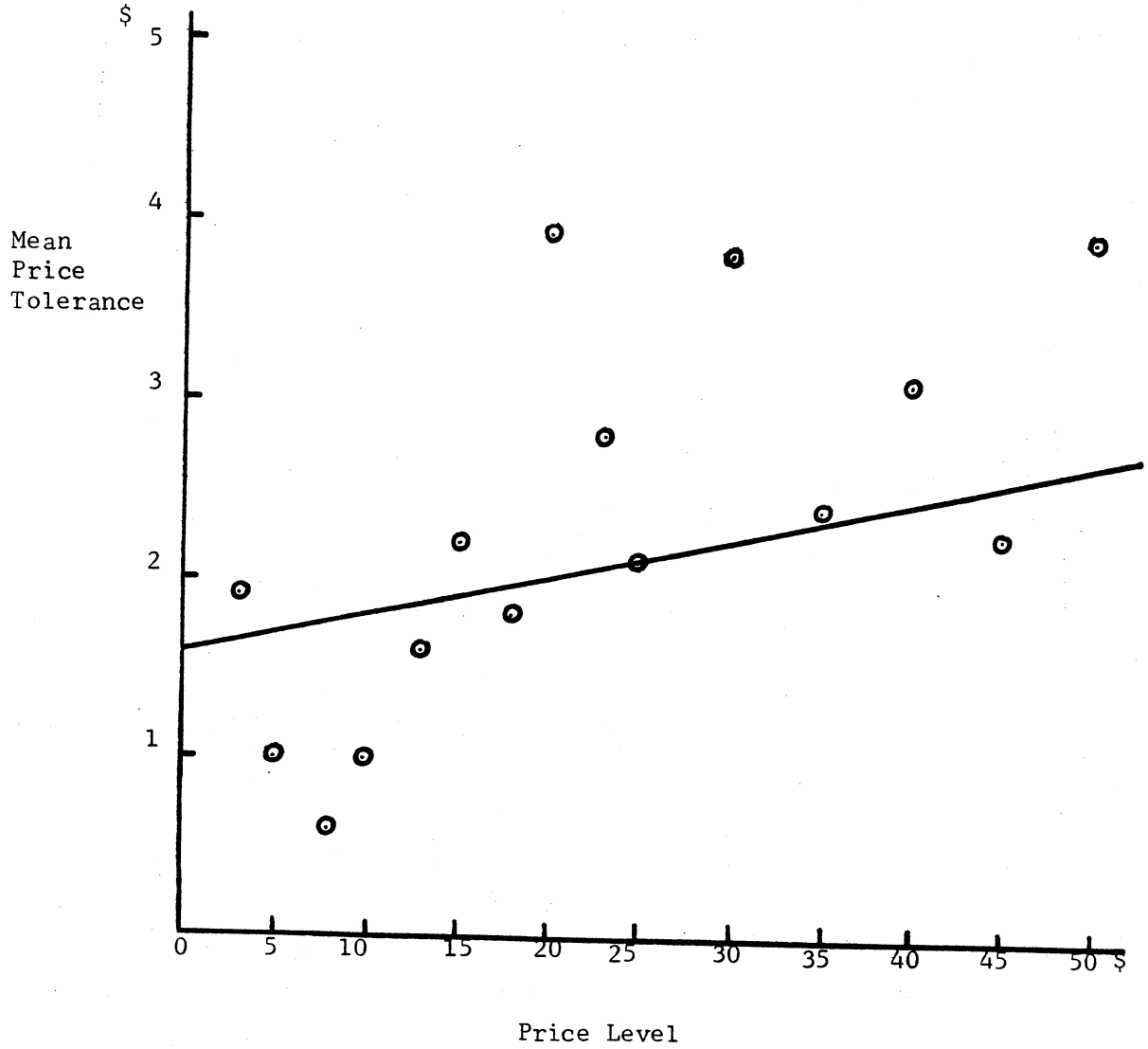


Figure 1. Mean price tolerance versus price level.

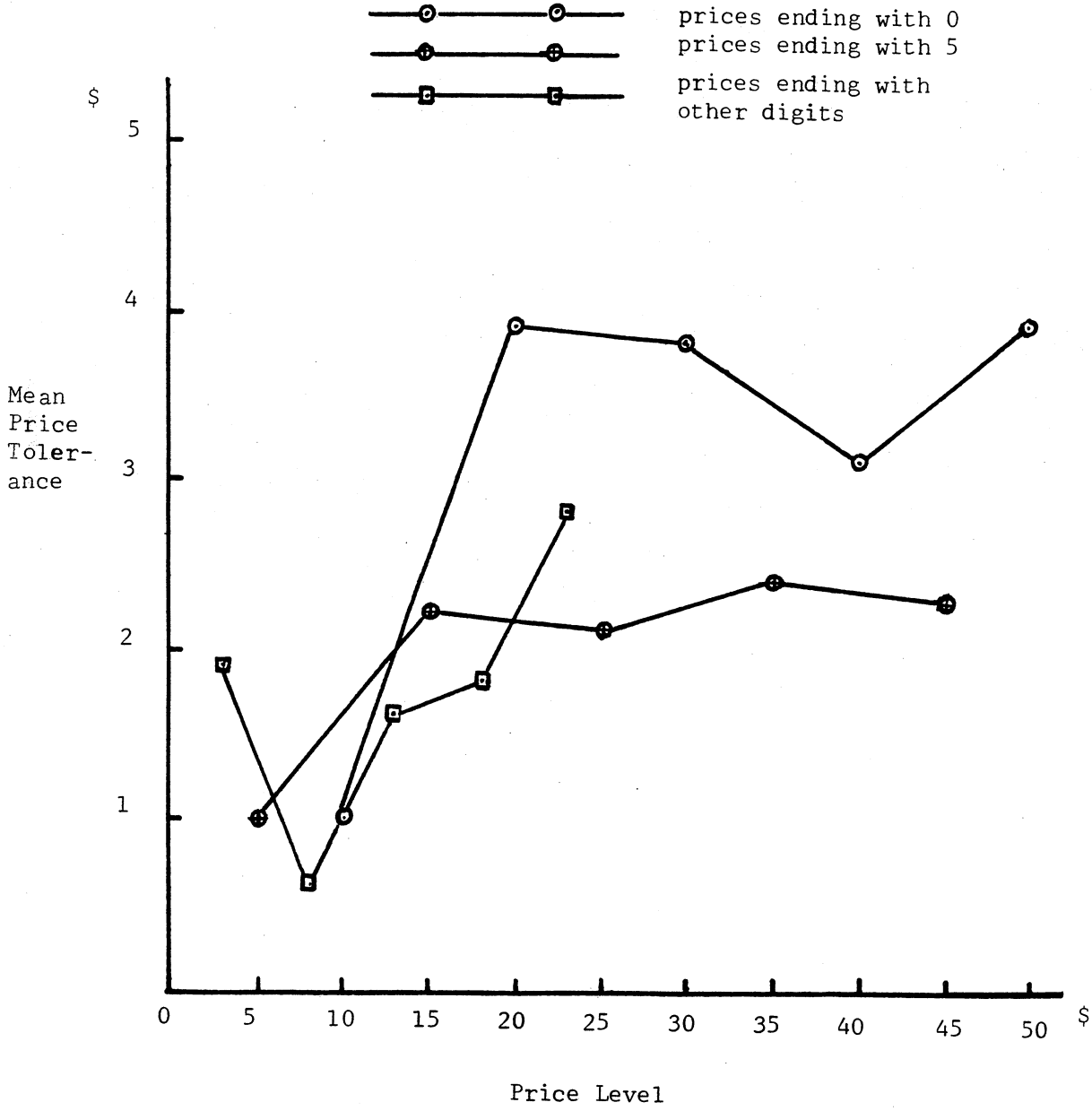


Figure 2. Mean price tolerance vs. price level broken into three plots.

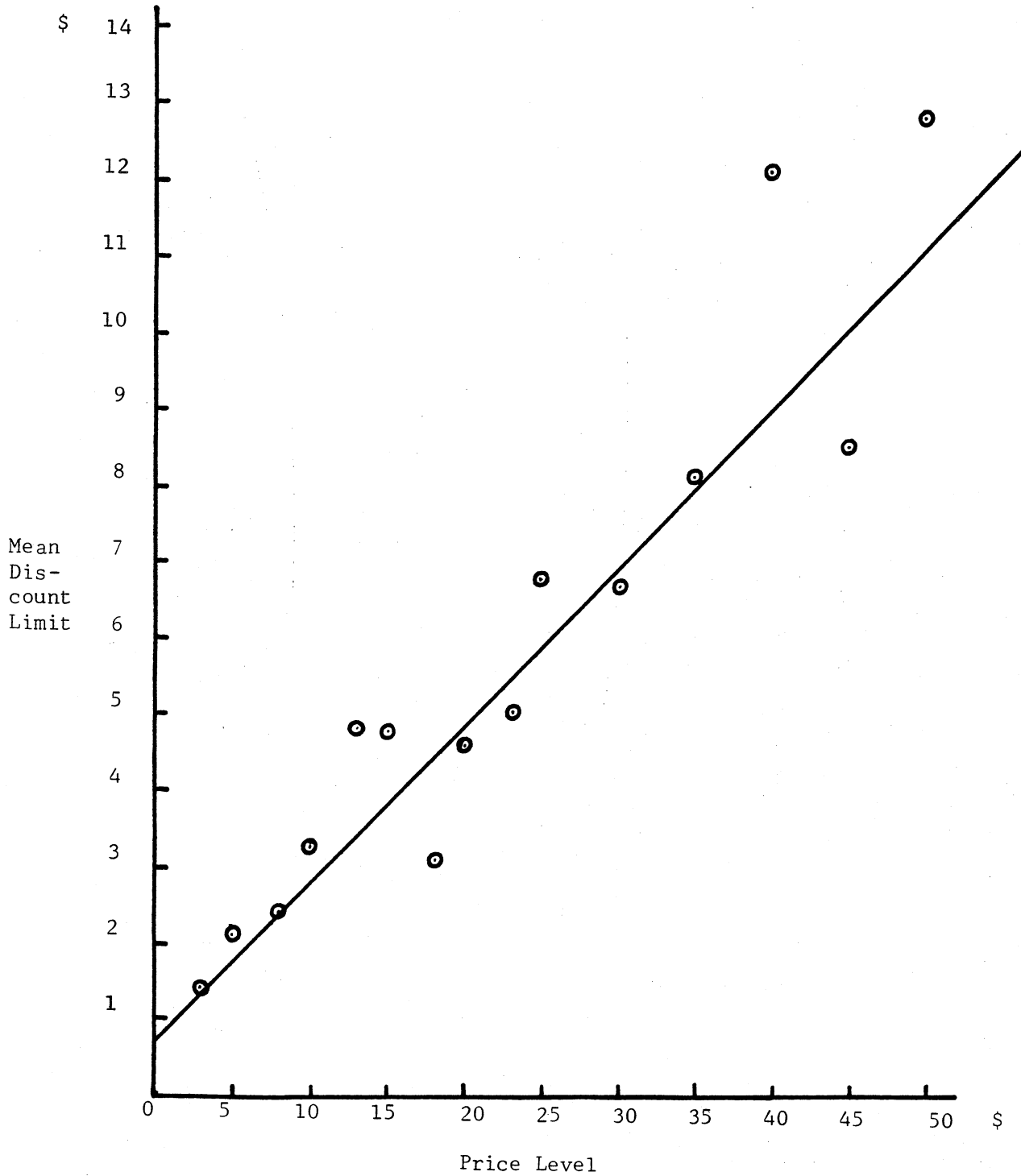


Figure 3. Mean discount limit versus price level.



$$Y = 1.53 + 0.04P \quad (1)$$

where Y is mean price tolerance and P is price level. The slope coefficient was positive and significant ( $p < 0.05$ ). Proportion of variance explained (R-squared) was 0.34 giving a correlation coefficient of  $R = 0.58$ . Thus, the hypothesized positive correlation between price tolerance and price level was weakly supported.

On closer inspection of the plotted points in Figure 1, it appears that price tolerance for price level ending in the digit 0, 5, or some other digit, follow different patterns. This led to drawing three plots as shown in Figure 2. Except for the price level of \$3, price tolerance for prices not ending with the digits 0 or 5 appears to follow a linear trend as price increases up to  $P = \$23$ . For price levels ending with the digit 0, price tolerance rises from \$1.0 for  $p = \$10$  and in the price range \$20 to \$50 shows little variation, staying roughly within \$3 - \$4. For price levels ending with the digit 5, price tolerance increases from \$1.0 for  $P = \$5$  and stays fairly flat at a value roughly \$2.25 in the price range \$15 to \$45.

Apparently, the structure of the number system is a factor to be included in explaining the response of the subjects to price changes.

A graph of mean discount limit versus price level is shown in Figure 3. There is less scatter of points and a more definite linear trend than in Figure 1. The linear regression equation obtained was:

$$Y = 0.76 + 0.22P \quad (2)$$

where Y is mean discount limit and P is price level. The slope coefficient was positive and significant ( $p < 0.01$ ). Proportion of variance explained (R-squared) was 0.89, giving a correlation coefficient of  $R = 0.94$ . Thus, the hypothesized positive correlation between discount limit and price level was confirmed. The high R-squared, confirmed by a significant F-test of the ANOVA, indicates that price level alone is highly effective in predicting discount limit.

Unlike the case of price tolerance, no clear cut patterns of discount limits are revealed when they are grouped according to the ending digit of the price level.

Hypothesis 3 states that mean price tolerance is less than mean discount limit at each price level. Columns 2 and 4 of Table 1 show that this is true except for  $P = \$3$  in which the relationship is reversed. Significance of the difference of means was tested at each price level. In ten of the fifteen cases t-values were significant beyond the 0.05 level, two cases were significant beyond the 0.10 level, and three cases (at \$3, \$20, and \$30) were not significant. Thus, there is support for the notion that buyers are more sensitive to price increases than to price decreases.

#### Discussion and Implications

It is clear that the highest price increase that the subjects would accept without changing their choices (price tolerance) is not predictable by price level alone. It would appear that when brand names are suppressed (as in this experiment) and price is responded to as a stimulus, that the structure of the number system as well as the perceived meaning of price strongly affect subjects' responses. Thus, price appears to be

multidimensional. The results of this study suggest that: (1) buyers are likely to accept bigger price increases for prices ending with the digit 0 than for prices ending with the digit 5, and (2) variation in the price tolerance for prices ending in the digits 0 and 5 is less than variation of price tolerance for price levels ending in other digits. This latter point perhaps reflects the greater ease of making mental computations with numbers ending in 0 and 5 than with numbers ending in other digits.

The above evidence suggests that for product classes where prices are relatively unstable, pricing a product at \$19.95 or \$19.99, for example, might lead to greater resistance to a future price increase than if the initial price had been \$20.

The finding that price tolerance is not zero, reflecting a reaction threshold on the part of the buyer, suggests that a seller could get away with some price increase without serious consideration of the competition or the short run effects on demand. To the extent that competing product alternatives in a product class are differentiated the seller has even greater latitude in raising price.

Price level predicted discount limit quite well. This indicates that other factors are relatively less important when a buyer reacts to a discount. The finding in this study that expected maximum discount was generally over 20 percent may depend upon the product class (here shoes). It is plausible that for some products, e.g., luxury items, the buyer does not expect any appreciable mark down, and may even doubt the quality of the product if the price is slashed.

In addition, a seller's fairness or integrity in pricing may be in question, if he applies a discount higher than what buyers expect. Exceptions may be when seasonal, perishable, or obsolete stock must be cleared. Thus knowledge of buyer's average discount limit may provide an approximate discount ceiling.

Some comments on methodology are in order. First, possible limitations in the execution of the study are the rather small size ( $n = 6$ ) of the experimental groups, and the mixing of male and female subjects in some groups. Women's shoes are generally priced lower than men's shoes and there might be sex differences in response to price changes. Fifty-two of the 180 students were female, but they were not evenly assigned to the thirty groups used. The data for those groups in which both sexes were equally represented were re-examined. There were three such groups for price increases and four groups for price decreases.

Mean price tolerances were significantly different in two cases out of three-- higher for males in one case and higher for females in the other (t-tests). Mean discount limits for females were higher than for males in all four cases of price decreases, but only one case was significant. The small sample sizes cast doubt on these tests. On the whole, sex differences do not appear to have been a major factor in the results.

Another comment on methodology is that if data on price tolerance had been obtained beyond \$23 for prices not ending with the digits '0' or '5', it would have been possible to see if the apparent linear increase of price tolerance with price (shown in Figure 2) would continue. Thus, hypothesis

1 might have been confirmed for prices not ending with the digits '0' and '5'.

Before the findings of this research can be generalized, there is need to replicate the study using non-student subjects such as housewives of a given socio-economic class, and trying other product classes.

#### Summary

The concepts of price tolerance and discount limit were introduced and used to study subjects' reactions to price changes. Price tolerance was defined as the largest increase in the price of a particular product offer that will make a buyer still consider making the purchase. Discount limit was defined as the largest decrease in the price of a particular product offer that will make a buyer still believe the original price. Mean price tolerance and mean discount limit were obtained for fifteen price levels for shoes. Percentage price tolerance ranged from 64 percent to 5 percent and mean discount limit ranged from 47 percent to 17 percent.

The hypothesis that price tolerance increases linearly with price level was weakly supported. An interesting effect indicating the multidimensional nature of price was revealed in the data. When brand images are suppressed or weak and price is responded to as a stimulus, it appears that the size of price increase that buyers will accept is determined as much by the number system (the ending digit of the price) as by the economic meaning of price.

The hypothesis that discount limit increases linearly with price levels received strong support. In 12 out of 15 price levels, mean price tolerance was significantly less than mean discount limit, providing new evidence that buyers are more sensitive to price increases than to price decreases.

#### FOOTNOTES

1. Nonyelu G. Nwokoye is a doctoral candidate in Marketing at the University of Massachusetts, Amherst. His thanks go to Kent B. Monroe for his valuable counsel in all phases of the study and to Alan G. Sawyer and David Finn for their assistance in administering the experiment.

#### REFERENCES

- Gabor, A., Granger, C., & Sowler, A. Comments on "psychophysics of prices." Journal of Marketing Research, May 1971, 8, 251-252.
- Kamen, J., & Toman, R. "Psychophysics of prices": A reaffirmation. Journal of Marketing Research, May 1971, 8, 252-257.
- Monroe, K. B. "Psychophysics of prices": A reappraisal. Journal of Marketing Research, May 1971, 8, 248-250.
- Pessemier, E. An experimental method for estimating demand. Journal of Business, October 1960, 33, 373-383.

Stapel, J. "Fair" or "psychological" pricing? Journal of Marketing Research, February 1972, 9, 109-110.

Uhl, J. Consumer perception of retail food price changes. Paper presented at the First Annual Meeting of the Association for Consumer Research, Amherst, Massachusetts, August 1970.



## INCORPORATING TESTS OF DATA AND MODEL VALIDITY IN COMMERCIAL AND ACADEMIC RESEARCH

Edgar A. Pessemier<sup>1</sup>

The important types of validity are reviewed as they apply to research in consumer and market behavior. Subsequently, two examples are presented, one from commercial research and the other from academic research. The commercial case emphasizes problem definition and measurement selection. The academic case focuses on the problem of developing a complex model where both the validity of the model and the measurements may be questioned. These two contexts serve to illustrate the need for early consideration of validity issues, incorporating validity tests in the research effort and reporting the validity of the measures and models along with other research results.

In the applied and academic realm, data collection and model selection are purposeful activities. Variables are designed to measure the properties the analysts wants to measure and a model is selected to expose the relations of interest. The validity of the data and model must be judged together with reference to explicit criteria. An incorrect choice of variables and measurement methods or the selection of an inappropriate model frequently yield faulty predictions.

### Validity and Reliability

Reliability is concerned with the consistency and accuracy of the measurement. Unreliable measures tend to mask significant relations and increase the quantity of data required to achieve useful results. The importance of reliability has been widely recognized and a large body of techniques have been developed to assist the research worker [Lansing and Morgan, 1971; Pessemier, 1971]. Although the subject of reliability is not central to the present discussion, a lack of reliability may indicate the absence of validity. For example, rushing subject responses or asking questions that are beyond the capacity of subjects to respond can result in low reliability and validity.

Validity deals with the meaning of measurement. The development of a consistent, precise measure does not guarantee its usefulness in any specific research context. If measurements do not measure what they are intended to measure, then the association of predictor variables to the dependent variables will be weak or misleading. But how can the validity of a measure be examined? From the point of view of consumer and market research, most emphasis is on predicting the future or present level of criteria from separately measured predictors. Predictive validity is tested by examining the capacity of data and a model to forecast future states of a dependent variable; will this month's consumer optimism index or sum score predict next month's automobile purchases. The score's construct validity may be demonstrated by reference to attitude theory. This theory indicates the measured level of an attitude (consumer optimism) should be associated with subsequent behavior (purchase of an automobile). In addition, the score may have an appropriate degree of

content validity if its components effectively represent the domain of consumer optimism. Given satisfactory evidence of content and construct validity, the predictive capacity of the optimism score specifies the measure's validity for a specific purpose, predicting the purchase of an automobile.

Concurrent validity measures the association between a predictor(s) and dependent measure(s) obtained at the same time; will the current occupation of the head of a household predict his (her) consumer optimism score. Here the case for construct validity is weak since it is harder to find strong theoretical grounds for believing present occupation will predict the current level of a consumer's optimism score. Also note that the two consumer optimism illustrations point up the need to examine model validity. The predictor score was obtained by simply summing several measures of consumer optimism. Even if an analyst boldly assumes all the essential variables were included, the use of equal weights may be challenged. A factor analysis of a larger set of candidate variables could alter both the variables employed and the method used to compute a score. Finally, note that no direct attention has been given to the question of how each basic variable was defined.

Although the measurement of complex constructs raise the most serious validity questions, widely used variables such as income and occupation may also be troublesome [Lansing and Morgan, 1971]. Income can be defined in many ways, each more or less relevant to various criterion, and occupation classes are largely arbitrary unless the purpose for which they are defined has been specified. Even the purchase of an automobile raises measurement issues. An automobile may be purchased new or used, with or without trade-in for cash or credit. Should these elements be considered and how should gifts and long-term rentals be handled?

#### Data Problem

In many types of scholarly work, the selection of measurements and models is made by the researcher. In commercial research, the analyst is often denied this luxury. If data comes from a syndicated service, the variables included are frequently limited by the needs of many users, the desire for comparability through time and the cost of data collection. In addition, the commercial analyst may have no capacity to influence the sample population from which data were collected or to vary the extent to which sample and response bias influence validity of the data for the roles it must play in his models. Furthermore, he may experience difficulty in comparing findings based on separate data bases. A company analyst studying the purchase of major appliances could use syndicated Trendex data, data from a proprietary consumer panel and consumer survey results published by the University of Michigan's Survey Research Center. Sample populations, time periods, methods of data collection, the coding of variables and applicable methods of data reduction may all potentially inhibit his capacity to make valid comparisons. In this illustration, however, sample sizes tend to be large and well defined, and many variables are sufficiently standard to permit a variety of cross data-base comparisons.

The commercial researcher's most important problem is not simply employing valid data and models to discover significant relations. His task is more difficult. He must find valid relations that are managerially important. That means the relations must be strong and the dependent variables must have useful economic meaning and suggest effective action. Clearly, the most important criteria to marketing decision-makers concern the purchase of a particular type of product or service, of the brands within a product class. Small differences in purchase rates or brand shares among various classes of

buyers may have large managerial significance. Since these differences can occur for many reasons, it is important that the predictor variables be valid and strongly associated with these criteria. It is equally important that tests of significance be purposefully selected. As Bass, Tigert and Lonsdale (1968) have observed, testing the worth of segmentation variables by referring to measures of explained variance may greatly understate their managerial value. The manager's first concern is group differences and the power to predict individual behavior may be of little or no interest.

Continuing to develop the appliance illustration, it is easy to think of a variety of strategic moves that could be associated with demographic predictors [Bratt, 1972]. They can assist product planners find market segments that are attractive targets for an innovative product or that are becoming stable replacement markets for mature products. In both cases, product life-cycle theory tells him that product design, pricing, and other considerations hinge on identifying the changing demographic characteristics of first time and replacement purchases [Smallwood, 1973]. Adoption and diffusion theory also suggests that buyer profiles change in predictable ways during the product's life-cycle [Rogers and Schoemaker, 1971]. For these reasons, a planner or analyst may attempt to measure the demographic properties of buyers and owners at various stages in the adoption cycle. A classification analysis of this type of data identified the following differences associated with income:

The Percentage of Owners in Each Income Class  
(10,250 Subjects)

Income	Owner Product A	Owner Product B	Owner Product C	Percent Sample
- 0 - to 8,000	11.0	27.0	20.0	38.0
8,000 to 14,999	42.0	44.0	47.0	40.0
15,000 +	47.0	29.0	33.0	22.0
Total	100.0	100.0	100.0	100.0
Percent Owners in Sample	24.0	33.0	54.0	

The association between income and ownership is extremely significant and managerially important but the expected higher proportion of high income to low income buyers/owners during the early phase of a product's life cycle cannot be clearly established from the data at hand. Product Type A is owned by 24% of the subjects and has a higher percentage of owners among high income groups than the other two products. On the other hand, Product B (33% owners) has a lower percentage of owners among the high income group than does Product C (54% owners). Also, more of Product Type B's buyers come from the low income group than do Product A's or Product C's buyers. Possible explanations abound. Is the sample biased, has the presence of a substantial number



of non-respondents to the income question rendered that variable invalid, are variables properly coded, is the theory strongly conditional on product type or is the theory faulty?

The analyst could examine his data from a variety of points of view, eliminating several possible sources of the observed "difficulty" by examining alternate studies to test for sample bias, comparing the non-respondents to the income question with respondents and repeating the analysis with a recoded income variable. In addition, the theory can be examined with the aid of other variables such as age and education. If the latter effort also produced "unexpected" results across the three product types, then attention should turn to ancillary measures and the statement of the research question.

An explanation of the observed results is simple to find and perhaps too obvious to have generated the initial concern. However, information is required that is not contained in the original data and cannot be directly measured. There is strong reason to believe that Product Type B is much closer to reaching the limits of its potential penetration than either Product Type A or C. In other words, the relevant populations are not common across the three product types. When percentage of ownership data were converted to percentage penetration, and the columns of the tables were rearranged, the theory was strongly supported across the predictor variables examined by the analyst.

The above illustration raised a common set of validity issues facing the commercial researcher. He must use the data at hand to investigate managerially important questions. A variety of methods can and should be used to establish the validity of his basic variables. The appropriateness of his analytical models should be examined in the light of the properties of these measurements. Finally, he must consider carefully the purpose of the analysis and satisfy himself that the research question has been correctly framed.

#### The Model Problem

A somewhat more difficult problem arises when the researcher is concerned with the validity of a model of consumer or market behavior. Often these models include high degrees of abstraction or make assumptions about processes that are not observable. For example, the models of Carroll, 1972 and Kruskal 1964, convert consumer perceptual judgments into spacial representations of objects and subjects. The spacial proximity of choice objects have meaning in terms of products' characteristics and consumers' power to discriminate among the competing alternatives. In addition, consumers' ideal locations can be placed in the same space, allowing the analyst to appraise how product characteristics are related to product preferences. Typically such joint space analyses have been based on product spaces obtained by decomposing similarity judgments. Recently, Pessemier [1973; Pessemier and Jones, 1974] has adopted the discriminant model to provide an alternative method of directly deriving a single consumer's product space from judged attribute levels for each choice object. This model is directly competitive with traditional multi-dimensional scaling procedures.

At least two validity questions surround the single-subject discriminant model. Data is required about the possible correlation of attribute levels within choice-objects. For example, if a consumer learns the safety of an automobile has been increased, in what direction and to what degree would this influence his judged level of another attribute, say cost? Data of this type can be collected in a variety of ways. The validity of each method is

hard to measure because the observations are individual judgments. Of course, the same problem has troubled researchers using similarity judgments.

The validity question in such cases can be examined in several ways. First, objects with known properties can be used and the congruence of judgment about them can be examined in relation to the objective measures. Second, judges' self reported experience in making the judgments can be appraised. Third, the stability and reliability of the judgments can be examined over time and across judges who have demonstrated a capacity to respond effectively to well understood perceptual tasks. Finally, the data can be employed in the discriminant model to develop perceptual product spaces and the performance of the spacial representations can be compared to product maps developed by standard scaling methods that use similarity judgments.

The flow diagram from a text book study [Jones and Pessemer, 1974] which used the above approach will clarify the nature of the validity tests:

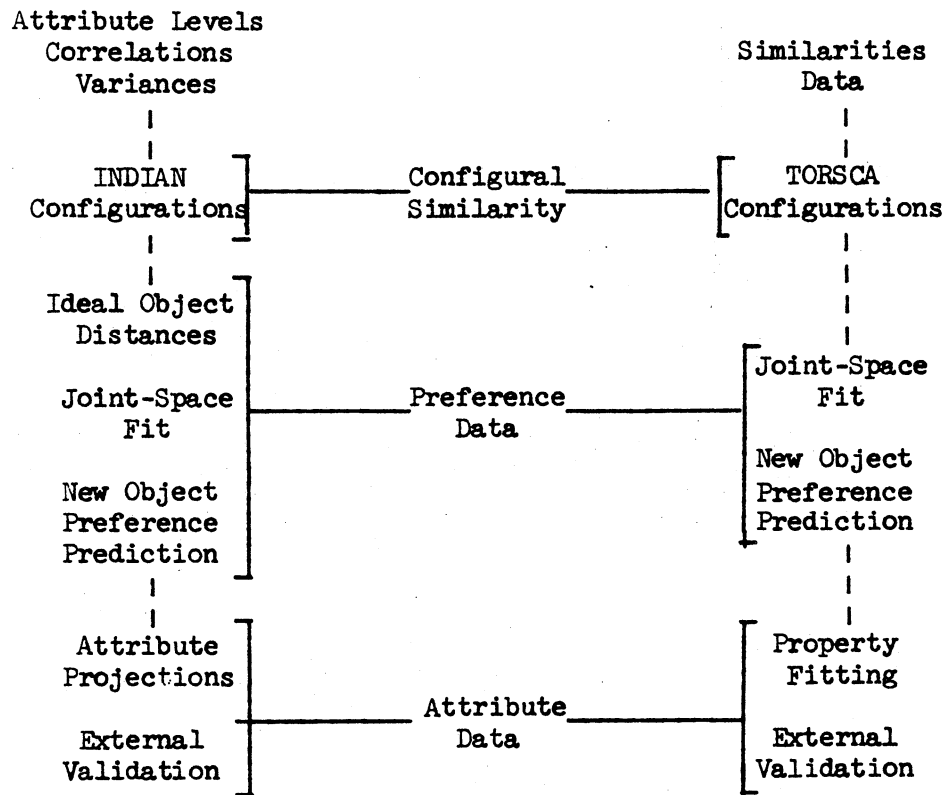


Figure 1. Analysis of configurations.

In the top phase of the analysis, two distinct classes of measurements were used, each in conjunction with an appropriate model and computer programs. Next, the congruence of these separately generated spacial configurations of text books was examined analytically. The reliability of the basic input data

was examined directly but the validity of the similarity judgments on one hand and the attribute level and correlational judgments on the other hand could not be directly studied. However, a comparative examination was made of the attribute data collected by different instruments. The configurations developed from these alternative data sources were also examined.

At the next level of analysis, the predictive qualities of the configurations were tested in joint-space. Prior to this test, however, the reliability and validity of several preference measures were studied to insure the integrity of the affective components of the joint-space model. With this assurance in hand, it was possible to use both types of configurations to predict the cognitive and affective placement of a "new" text book not used to develop the original product space. This form of predictive testing is especially helpful in establishing the validity of the research procedures.

Finally, the object and attribute projections in the discriminant space were compared with separately fitted attributes in the space derived from similarity judgments. This step helped establish the cognitive compatibility of the two procedures and indicated how well the directly judged attributes spanned the space generated from similarity judgments.

Under certain limiting conditions, the single subject discriminant model can be shown to produce the same space as a scaling model. This mathematical property is of little real comfort since the required conditions cannot be easily met. Even if they were, the validity of a configuration in the context of the joint-space model in which it is most frequently used cannot be directly established. Therefore, the models have to be judged in relation to the levels of content and predictive validity that can be demonstrated in specific research tasks. In this regard, four research efforts have been launched to study the single-subject discriminant model's comparative validity across four diverse types of choice objects; textbooks, beer, recreational programs and nuclear power facilities [Jones and Pessemier, 1974; Moore, 1974; Jones, 1974; and Braden, 1974]. The findings produced by the above procedures and similar comparisons made in the other studies cited offer considerable support for the new methodology. As the superiority of the one or another of the models is demonstrated by repeated applications, increased reliance may be placed in the model and particular type of data which it employs. Successful experience with a growing range of research contexts should increase the analyst's confidence.

#### Conclusion

The characteristics and importance of validity has been outlined as it appears in several applied and theoretical realms. The scholar has more freedom to specify the nature of his measurements and the structure of his models. On the other hand, an applied worker will typically work with larger samples and simpler, better understood measurements and analytical methods. Each type of investigator has an obligation to continuously monitor the validity of his research procedures. If analysts remain sensitive to the problem, numerous opportunities arise to test the validity of measurements and models. In doing so, it is imperative that the results be reported. A model for the scholarly community can be found in the work of the Michigan Institute for Social Research. Robinson and Shaver, (1969), have documented the measurement and validity properties of a wide range of instruments in the field of social psychological attitudes. Lansing and Morgan present a useful general discussion of the appropriateness of the variables and data reduction methods employed

in IRS economic surveys. Finally, there is a growing scholarly literature. Under the sponsorship of the National Bureau of Economic Research, the Annals of Economic and Social Measurement regularly publishes works on measurement of interest to the consumer and marketing research community.

In the non-academic research community, the Census Bureau (1970), and other government agencies have shown increasing concern with validity. Considerable attention is given to improving the data collection and reporting the methods employed. Also, firms producing a wide range of commercial research data are finding their buyers have become more sophisticated, calling for relatively complete disclosure about the quality of data, [Pessemier, 1971; Pessemier and Bruno, 1972]. The work of various professional associations has contributed to these developments. Clearly, the academic and applied research communities should continue the healthy trend towards improving validity standards in data collection and model development.

#### FOOTNOTE

1. Edgar A. Pessemier is Professor of Industrial Administration, Krannert Graduate School of Industrial Administration, Purdue University.

#### REFERENCES

- Bass, F. M., Tigert, D. J., & Lonsdale, R. T. Market segmentation: Group versus individual behavior. Journal of Marketing Research, 1968, 5, 264-270.
- Braden, P. A study of attitudes toward nuclear power generating facilities. Unpublished study in progress, Bureau of Business Research, University of Michigan, August 1974.
- Carrol, J. D. Individual differences and multidimensional scaling, In R. N. Shepard, A.K. Kimball and S.B. Nerlove (Eds.), Multi-dimensional scaling. Vol. 1. Theory. New York: Seminar Press, 1972, 105-157.
- Jones, W. H. & Pessemier, E. A. Single subject discriminant configuration: An examination of reliability and validity and joint-space implications. Institute Paper No. 451, Krannert Graduate School of Industrial Administration, Purdue University, March 1974.
- Jones, W. H. A multidimensional joint-space approach to the design and selection of outdoor recreation facilities. Unpublished doctoral dissertation in progress, Purdue University, 1974.
- Kruskal, J.P. Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis. Psychometrika, 1964, 29, 1-27.
- Lansing, J. B. & Morgan, J. N. Economic survey methods Institute for Social Research, University of Michigan, May 1971.
- Moore, W. L. 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, August 1974.

- Pessemier, E. A. Data quality in marketing information systems. Institute Paper No. 297, Krannert Graduate School of Industrial Administration, Purdue University, February 1971.
- Pessemier, E. A. Single subject discriminant configurations. Institute Paper No. 406, Krannert Graduate School of Industrial Administration, Purdue University, April 1973.
- Pessemier, E. A. and Bruno, A. V. An empirical investigation of the validity of selected attitude and activity measures. In M. Venkatesan (Ed.), Proceedings 3rd annual conference. Iowa City, Iowa: Association for Consumer Research, 1972.
- Pessemier, E. A., & Jones, W. H. Joint-space analysis of the structure of affect using single-subject discriminant configurations, Part 1. Institute Paper No. 435 (Rev.), Krannert Graduate School of Industrial Administration, Purdue University, February 1974.
- Pratt, R. W. Marketing applications of behavioral economics. In B. Strumpel, J. N. Morgan and E. Zahn (Eds.), Human behavior in economic affairs San Francisco: Jossey-Bass, 1972.
- Robinson, J. P. & Shaver, P. R. Measure of Social Psychological Attitudes, Institute for Social Research, University of Michigan, August 1969.
- Rogers, E. M. & Schoemaker, F. F. Communication of innovations. New York: The Free Press, 1971.
- Smallwood, J. E. The product life cycle: A key of strategic marketing planning. MSU Business Topics, Winter, 1973, 29-35.
- U. S. Bureau of Census Census bureau methodological research, August, 1970.

DETECTING DEMAND CHARACTERISTICS IN LABORATORY EXPERIMENTS<sup>1</sup>  
IN CONSUMER RESEARCH: THE CASE OF REPETITION-AFFECT RESEARCH<sup>1</sup>

Alan G. Sawyer<sup>2</sup>

University of Massachusetts

Demand characteristics in laboratory experiments are various aspects of the experimental environment which may be used by a subject as cues to the experimenter's hypothesis and as a guide to appropriate behavior. Four research methods--the non-experiment, the post-experimental inquiry, manipulation of hypothesized demand cues, and hetero-method replication--can help to detect demand characteristics. Research investigating the effects of repeated stimulus exposures on measures of acceptance, liking, or preference is presented to demonstrate the problems of demand bias in a particular repetition-affect experimental paradigm and to illustrate each of the four methods of detection.

The problem of demand characteristics in laboratory experiments concerns aspects of the experimental environment which may be used by the subject as cues to the experimenter's hypothesis and as a guide to appropriate behavior in the experiment.

Insofar as the subject cares about the outcome, his perception of his role and of the hypothesis being tested will become a significant determinant of his behavior. The cues which govern his perception - which communicate what the experimenter hopes to find - can therefore be crucial variables.... They include the scuttlebutt about the experiment, its setting, implicit and explicit instructions, the person of the experimenter, subtle cues provided by him, and, of particular importance, the experimental procedure itself. All of these cues are interpreted in the light of the subjects' past learning and experience. Although the explicit instructions are important, it appears that subtler cues from which the subject can draw covert or even unconscious inference may be still more powerful (Orne, 1969, p. 146).

The various types of demand characteristics and how they may affect subjects' behavior have been extensively discussed by social psychologists (e.g., Rosenthal and Rosnow, 1969; Weber and Cook, 1972; Rosnow and Aiken, 1973) and by consumer researchers (e.g., Olson, 1974; Sawyer, 1974). Four methods can help to detect demand characteristics. To both demonstrate the problems of demand characteristics and to illustrate the four methods, research investigating the effects of repeated stimulus exposures on measures of acceptance, liking, or preference for that stimulus will be discussed.

#### Repetition and Affect

Complete reviews of the effects of repetition on affective measures are available (e.g., Zajonc, 1968; Sawyer, in press) so only a brief summary of

various procedures and results will be presented here. Zajonc (1968) reported that increasing the number of exposures of such stimuli as nonsense syllables, Chinese characters, and Turkish adjectives in a series of experiments resulted in increased liking for those stimuli. Zajonc exposed slides of the stimuli for two seconds each and varied the exposure level from 0, 1, 2, 5, 10, to 25 and found a positive monotone relationship between the logarithm of frequency of exposure and liking as measured by the subjects' estimates of whether the foreign words meant something good or something bad. This relationship has since been replicated several times. One experiment also found that the frequency-affect relationship remained the same when measured one week later. Moreover, with the Chinese characters as stimuli, the effects of increased exposure upon rated goodness of their meaning remained positive at exposure levels as high as eighty-one (Zajonc et al., 1971).

Research into the effects of repeated exposures on affective responses has produced mixed results in more consumer-oriented investigations. Becknell, Wilson and Baird (1963) studied the effect of repeated nonsense syllables on "brand choice." Along with print advertisements and landscapes, a set of nonsense syllables with frequencies of one, four, seven, and ten were shown. After one showing of the set, each respondent was asked to choose a gift of nylon stockings from two brands marked by nonsense syllables seen in the presentation. A second showing was followed by a paper-and-pencil ranking of four "brands." Responses to both measures were positively affected by repetition. Miller, Mazis, and Wright (1971) found positive effects of 0, 2, 10, and 20 exposures of different nonsense syllable brand names on rated liking for different characteristics linked to the brand names after they had been repeated. However, no effect of repetition on liking was found when subjects were also able to actually sample the product with which the repeated brand name was linked.

Consumer studies with actual product advertisements have generally found either neutral or negative affective responses to repetition. Light (1967) measured ad evaluation and ad satiation (the trend to neutrality on five semantic differential scales). Color slides of print advertisements were exposed in frequencies ranging from one to sixteen in two experimental conditions: high exposure density (fifty total exposures) and low density (seventy-five total exposures). Light found no evidence of a trend over repeated exposures to ad satiation and no effect of repetition on ad evaluation. Moreover, in the high density condition, there was a negative relationship between ad frequency and ad evaluation. The lack of any positive relationship of reactions to the ad and repetition in comparison to Zajonc's positive results is noteworthy.

Ray and Sawyer (1971) developed a laboratory procedure which measured several responses including brand evaluation (rating on a good-bad scale) and brand purchase intention (which brand in a product class would the subject be most likely to purchase). The true purpose of the experiment was disguised, and volunteer female shopper subjects were led to believe they were looking at a new shopping-by-television demonstration. Color slides of print ads were repeated from one to six times for three brands within several product classes. Although brand purchase intention was positively affected by repetition, there was no significant effect of repetition on brand evaluation. Moreover, when the relative rating of the brand in question was used to develop a score of preference to other brands in the product class, the effects of repetition coincided with a significant decrease in preference. With a different version of Sawyer's disguised laboratory procedure and with television as well as print advertisements as stimuli, Heeler (1972) also failed to find any posi-

tive increases in brand evaluation due to repeated ad exposures.

#### Demand Characteristics in Repetition Research

From the above review, it seems that positive effects of repetition on affect are more likely to result from experimental procedures in which the repeated exposures are presented with little or no accompanying "cover" or rationale for the subject and where no information other than the experimentally manipulated number of stimulus exposures is available to the subject. Such artificial settings may be antecedents of demand bias. Inconsistent past results might be explained by differences in the environment of the experiments and in the presence of demand characteristics which may cue a subject to respond positively to repeated stimuli.

One experimental paradigm that includes both of these demand-prone characteristics is the one used by Zajonc and his colleagues. An examination of the typical Zajonc experimental procedure from a subject's viewpoint reveals that his role in the experiment might be perceived as quite strange and artificial. The task of rating the foreign adjectives is a near impossible task. However, a subject realizes that the experimenter is performing the experiment for some purpose and expects the subject to rate the stimuli in some meaningful way. Thus, many subjects are likely to search for any available cues to differentiate the stimuli. Since the stimuli differ somewhat in appearance, some patterns of "goodness" may be discerned in these differences. However, the major difference in the stimuli is that, for some unexplained reason, the words were repeated several times but at different exposure frequencies. The fact that no warning or rationale for that repetition was offered by the experimenter adds to the artificiality of the experimental setting and may further focus the subject's attention to that variable. A subject who does not reject the task as too difficult or inane is likely to seize upon the widely manipulated number of exposures to differentiate his ratings on the stimuli. To the extent that the subject perceives that the experimenter expects some use of the exposure variable in the rating task and to the extent he believes the relationship should be positive, the subject may increase his ratings for stimuli that were exposed more often.

Such hypothesized demand cues might help to explain some of the inconsistent results between Zajonc's research and consumer research based on his work. Repeated consumer stimuli such as advertisements, actual products or brand names are presented in the context of the consumers' experiences and current perceptions and not in a vacuum where the number of exposures is the only available information. Also consumer research finding no positive effects of repetition on liking has often included deception or other explicit attempts to present the repeated stimuli in some natural or justified manner. Demand characteristics might account for the fact that brand evaluation or attitudes are often not affected by repeated ad exposures but measures of purchase intention are positively affected. Subtle demand bias might affect the simpler purchase intention response of simply checking a brand but might be not strong enough to significantly alter a subject's value structure of his perceptions of a brand or its characteristics.

There are four research modes that can examine the validity of such conjecture about demand characteristics. These include: 1) the non-experiment, 2) the post-experimental inquiry, 3) manipulation of suspected demand cues, and 4) hetero-method replication. Investigations of each of these types have recently focused upon the Zajonc paradigm. Details of new research and summaries of published research are presented both to expand on the potential de-



mand bias in repetition-affect research and to illustrate the use of these methods.

### The Non-Experiment

A non-experiment involves a reenactment of all experimental procedures except the actual treatment which is instead only simulated (Orne, 1969). New subjects from the same population sampled in the experiment in question are asked to role-play the subjects in the actual experiment. They then undergo the same experimental procedure as the actual subjects. They are shown the same room and the employed equipment, read the same instructions, and, after the experimental treatment is described, asked to produce data just as if they had actually undergone the experimental treatment. A comparison of results from the non-experiment and the actual experiment can offer insights into the role of hypothesized demand characteristics. If results are similar, the cause-and-effect relationship cited in the experiment is open to question, and further investigation is warranted. Although similar results do not prove the presence of bias due to demand characteristics, such results do demonstrate the plausibility of the alternative demand characteristics explanation that cues in the experimental procedure--instead of or in addition to the cited independent treatment variable--helped to produce the observed experimental result.

This writer designed a "non-experiment" that simulated the Zajonc procedure but did not actually repetitively expose any stimuli to the subjects. The experimental hypothesis was that there would be a significant effect of "exposure" even though no actual exposure occurred. Seventy-two subjects were told that they were going to participate in an experiment where they would be read only a description of an actual experiment and then would be asked to behave as if they were subjects in an experiment like the one described. They were then read the following description:

A year ago, a psychological experiment about the learning of a foreign language was performed at the University of Michigan. Here is what happened:

Subjects were seated at a table in front of a slide screen. After being told that they were participating in an experiment about the learning of a foreign language, they were asked to study a following series of slides. The room was darkened, and they were shown a series of 86 slides--each of which was exposed, one after another, for 2 seconds, or a total of almost 3 minutes. The slides were pictures of characters which looked like Chinese letters or words. During the slide presentation, it became apparent that many of the characters were being shown more than once. In fact, by the time the slides were over, some of the slides had been repeated as many as ten or twenty times.

We would like you to now sit and try to imagine sitting and watching these slides for almost 3 minutes--one slide every 2 seconds and several being repeated quite frequently.

After sitting for three minutes, the subjects were read Zajonc's explanation that the characters were Chinese symbols and that, as an estimate of their meaning, their rating of the "goodness" of the symbols was sufficient. In addition, they were given the following rating instructions:

Below are the characters which were shown in the series of slides. Since you did not actually see the slides, we have indicated, next to each word, the number of times that word was shown in the pre-

vious experiment. Note that the number of repeated exposures varied from 0 to 25. Please circle the point on the evaluation scale that represents your best guess of the degree of favorableness of the meaning of that foreign adjective word. (Note that "0" and "6" represent the two extremes of the scale and that "3" is the neutral scale point.)

The subjects then rated the twelve Chinese characters from the Zajonc experiment. After rating the characters, they were asked four questions about criteria used in the rating task and about the experimental hypothesis.

As in the Zajonc experiment being duplicated, the twelve characters used as stimuli were counterbalanced against six indicated frequencies (0, 1, 2, 5, 10, and 25), such that each character occurred equally often in each of the six frequencies, but for different subjects. Thus, for each subject, two characters were indicated as having been exposed for each frequency level, and each character appeared in each exposure condition to some subjects. In addition, to assess the importance of measurement "set," half of the subjects used scales marked "bad-good" and half used scales marked "good-bad." Six subjects per scale-character-frequency cell were used.

The experimental hypothesis was that subjects would increase the rated goodness of the characters as "frequency" increased from 0 to 25, and the results supported this hypothesis (see Figure 1). The rated goodness increased from 2.56 on a seven-point scale to 3.50 ( $F=6.21$ , 5 and 720 df,  $p < .001$ ). As in Zajonc's experiment, the ratings of eleven of the twelve characters were higher at 5, 10, and 25 exposures than at 0, 1, and 2 exposures. The fact that subjects who did not actually view the repeated stimuli but who did experience the experimental setting and accompanying cues produced results quite similar to those of subjects who did actually view the repeated exposures lends some support to a hypothesis that demand cues in the experimental setting and procedure might have helped to produce Zajonc's observed results.

#### The Post-Experimental Inquiry

Upon completion of a subject's participation in an experiment, it should be standard procedure to interview the subject to determine his reaction. Such a practice takes advantage of the fact that human subjects can talk and reflect upon their experience. This inquiry may include a debriefing which explains the purpose of the research, the subject's contribution, and the reason for the procedure including deception if used. Concerning demand bias, however, the prime purpose of the post-experimental interview is to determine any suspicions the subject may have developed about the purpose of the experiment, the subject role he might have adopted, and why he behaved as he did.

In the above non-experiment, answers to post-experimental questions designed to reveal subjects' reactions to the experiment were analyzed. When asked to describe any guidelines or criteria used in rating the characters, fourteen or almost 20% indicated that they used repetition to indicate goodness. The remainder indicated that they used only the appearance of the characters or used no criteria at all. An examination of the ratings of those who claimed to use repetition as a positive indicator revealed that these subjects accounted for a large portion of the overall increase with frequency. Table 1 shows the rated goodness of the repeated characters for the two groups. Those who claimed frequency as a criterion exhibited pronounced increases whereas the others showed only a slight increase. A post-hoc analysis of var-

Figure 1

Comparison of Frequency Effects in  
in Zajonc and "Non-Experiment"

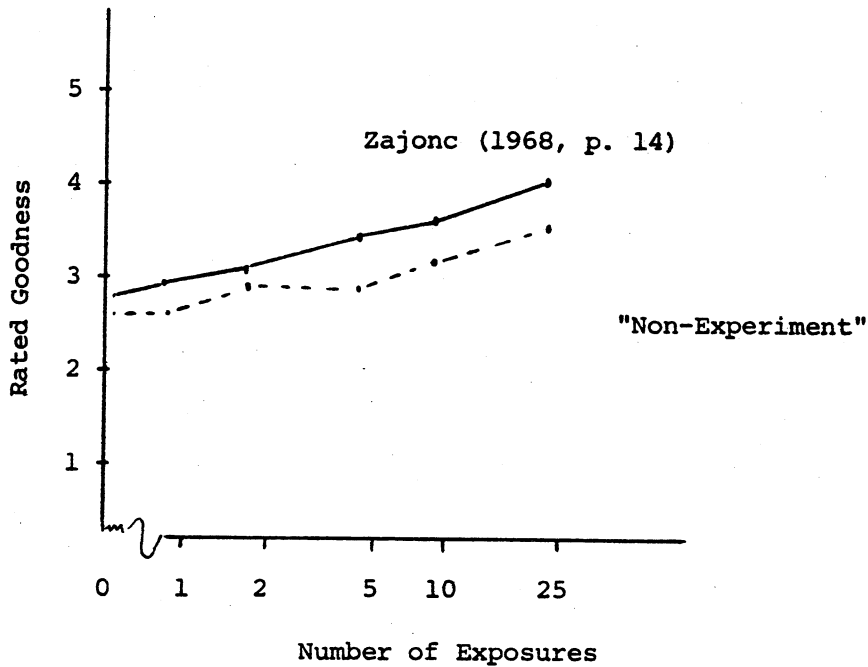


TABLE 1

Repetition-Affect Results of Subjects  
Claiming to Use Frequency as an Indicator of  
Goodness and of Those Who Did Not

Rating Criterion	Repetition Level					
	0	1	2	5	10	25
Used Repetition As a Criterion (N = 14)	1.57	1.86	2.28	2.50	3.00	5.14
Did Not Use Repetition (N = 58)	2.80	2.72	2.99	2.93	3.09	3.10
Total (N = 72)	2.56	2.55	2.85	2.85	3.07	3.50

iance with frequency and subjects' cited rating criterion as factors indicated a significant effect of both frequency ( $F = 5.87$ , 5 and 852 df,  $p < .001$ ) and its interaction with the cited criterion ( $F = 9.97$ , 5 and 852 df,  $p < .001$ ).

In answer to a question which asked the subjects to state what they believed to be the experimenter's hypothesis, eleven or 15% stated that more frequent exposures were expected to increase goodness. Seven others stated an op-

posite hypothesis; repetition would result in less goodness. The remainder of the subjects either paraphrased the opening instructions or stated some psychological hypothesis about the appearance of the characters. Table 2 shows how these subjects reacted to the indicated frequency of the rated char-

TABLE 2

Repetition-Affect Results of Subjects  
With Different Estimates of Experimenter's Hypothesis

<u>Estimated Hypothesis</u>	<u>Repetition Level</u>					
	<u>0</u>	<u>1</u>	<u>2</u>	<u>5</u>	<u>10</u>	<u>25</u>
Repetition = Good (N = 11)	2.77	2.27	2.46	2.91	3.05	4.27
Repetition = Bad (N = 7)	3.86	3.50	3.00	2.71	2.86	2.07
All others (N = 54)	2.35	2.48	2.91	2.86	3.10	3.53
<hr/> Total (N = 72)	2.56	2.55	2.85	2.85	3.07	3.50

acters. The first group demonstrated a large increase with repetition, and the second showed a decrease. The latter group, which did not mention frequency as part of the experimental hypothesis, showed an increase with greater frequency and nearly duplicated the reactions of all subjects averaged together. Both frequency ( $F = 5.61$ , 5 and 846 df,  $p < .001$ ) and its interaction with subjects' estimates of the experimenter's hypothesis ( $F = 1.80$ , 10 and 846 df,  $p < .06$ ) were significant. Thus, it appeared that subjects' reactions to increased frequency coincided with their estimates of what was expected of them by the experimenter - a result consistent with a demand characteristics explanation of Zajonc's positive relationship between repetition and effect.

Similar results of other post-experimental inquiries have been reported. Stang (1974) found very close relationships between subjects' intuition about the experimenter's hypothesis and their affective ratings of repeated Turkish words. Burgess and Sales (1971) found that subjects who reported feeling relatively positive about the experiment and their participation exhibited a significant positive relationship between the number of exposures and liking, whereas subjects who felt relatively negative about the experimental context exhibited no effect of exposure.

#### Manipulation of Demand Characteristics

When demand characteristics are suspected, experimental variation of the hypothesized cues can help to explore their effects. If the effects of the hypothesized demand cues are monotonic, then increases (or decreases) in the suspected aspects of the experimental environment should increase (or decrease) their effects.

For example, Suedfeld *et al.* (1971) manipulated the direction of measurement "set" as either positive or negative ("tell me the extent that this sym-

bol means something good [bad]"). They found that a good set coincided with positive monotonic effects of repetition on attitude whereas a negative set led to an inverted U-curve relationship. Burgess and Sales (1971) hypothesized that classical conditioning between positive evaluation of participation in the experiment and high frequency stimuli might account for Zajonc's positive results. In an experiment that manipulated frequency of nonsense words and the positiveness of paired-associated words, the rated liking of the nonsense words at different frequency levels varied with the context of the associated words. As hypothesized, a negative relationship was found for the negative context conditions, and a positive relationship was found for the positive conditions.

The relative effects of other potential demand cues in repetition experiments could be accomplished by manipulating the suspected procedural cues contrary to the direction of the claimed causal variable of repeated stimulus exposure. For example, subject's perceived personal evaluation from responding positively, negatively, or neutrally to repetition could be experimentally varied (e.g., Rosnow *et al.*, 1973). An alternative method would be to first expose subjects to the various levels of exposure and then subsequently mislead them by stating that low frequency stimuli were exposed at high levels and vice versa. A negative repetition-attitude result would indicate a greater influence of demand characteristics than the actual amount of exposure. This latter method would require that the results of the actual repetition conditions remain at the time of the deception and that subjects would forget the real repetition levels and believe the experimenter's false repetition "reminder."

#### Heteromethod Replication

Multiple operationalism of independent variables and accompanying procedures can help to refute questions about confounding demand cues. If more than one procedure yielded similar results, the result (and accompanying theoretical foundation) would obviously carry more confidence.

Repetition experiments with different procedures or designs have both refuted and supported a demand characteristics explanation. Saegert, Swap, and Zajonc (1973) studied the effects of repeated personal encounters on interpersonal attractions and manipulated the pleasantness of the experimental context by presenting pleasant and noxious odors or liquids to smell or taste. Contrary to the results of Burgess and Sales, positive effects of frequency were found in both pleasant and unpleasant contexts. In addition, Zajonc and Rajecki (1969) tested the effects on liking of nonsense words when exposure was experimentally varied from one to twenty-five by means of newspaper inserts over a month's time. Randomly selected readers of the newspaper were interviewed and a positive relationship between word exposure and rated goodness of meaning was found. Although the field results are impressive, field replication does not preclude the possibility of demand bias (e.g., Rosen, 1970). A post-experimental interview probing why the more heavily repeated words were rated more positive would have been very helpful.

Moreover, some of Zajonc's own lab experiments with variations in the number of exposures and experimental design have offered additional support for an alternative demand bias explanation. Zajonc, Swap, Harrison, and Roberts (1971) reported that the absolute number of exposures appeared to be much less important than the relative number. The results of an experiment which tested

0, 1, 3, and 9 exposures did not differ significantly from those of experiments with exposure levels of 0, 1, 3, 9, and 27 and 0, 3, 9, 27, and 81. Such results would be expected if past effects of repetition were actually due to demand bias whereby repetition was perceived as merely a cue to differentiate ratings as expected by the experimenter. Furthermore, no significant effect of exposure of initially unfamiliar characters was found in a fourth experiment in which the manipulated exposure treatment was across-subjects rather than the typical within-subjects design. That is, no effect of frequency was obtained when subjects were not able to compare different levels of exposure--a design that could not have artificially highlighted repetition as a cue to discriminate ratings of the tested stimuli.

#### Conclusion

One objective of this paper was to present research evidence of demand bias in a particular repetition-affect experimental paradigm. It seems safe to conclude that the effects of repeated exposures are highly sensitive to small changes in the experimental procedures and environment. However, the case for demand characteristics as a cause of a positive monotone repetition-affect relationship instead of the cited independent variable of the number of exposures has not been conclusively proven. Supporting evidence comes from the facts that a minority of subjects who use a positive connotation of repetition as a response criterion account for a disproportionately high amount of the observed positive relationship, subjects' reactions to repetition coincide highly with their estimates of the experimenter's hypothesis, manipulations decreasing either the positiveness of the experimental context or of items associated with the repeated stimuli can eliminate or even reverse the positive repetition-affect relationship, and positive repetition-affect results depend on the subjects' being able to compare relative exposure levels instead of absolute levels.

However, if subjects do interpret and react to their estimates of the experimenter's hypothesis, none of the demand bias research has provided an adequate explanation of why subjects tend to perceive a hypothesized repetition - good relationship instead of a negative one. Some conjecture might involve the tendency to adopt a "good" subject role (Orne, 1969) or the overall positiveness associated with science and one's participation in an experiment (Burgess and Sales, 1971). The fact that experiments with advertisements as repeated stimuli are more apt than subjects in more sterile experimental contexts to yield neutral or negative effects of repetition might be due to subjects who negatively evaluate the social and aesthetic values of advertising (Bauer and Greyser, 1968) and who are more apt to adopt a "negative" or "apprehensive" subject role (Weber and Cook, 1972) or to feel negative about their participation in the experiment. Currently, however, there is no evidence to support such hypothesized subject negativism.

The second goal of this paper was to illustrate the use of four research methods to evaluate the possible presence of demand characteristics in a given understanding of the basic process under study as well as improved experimental methodology. Other than repetition, many areas in consumer research would benefit from similar demand characteristic investigations. These areas might include quality connotations of price, conformity in consumer choice behavior, bargaining behavior between buyers and sellers, and post-purchase dissonance (Sawyer, 1974). More careful experimental designs and procedures could help to reduce demand bias. More consumer research should con-

centrate on hetero-method replications and manipulations of suspected confounding demand characteristics, and the use of non-experiments as pretests for planned experiments and the inclusion of post-experimental inquiries which probe for demand bias should be standard experimental procedures.

#### FOOTNOTES

1. This research was supported by a grant from the Research Council of the University of Massachusetts.
2. Alan G. Sawyer is Assistant Professor of Marketing at the University of Massachusetts at Amherst.

#### REFERENCES

- Bauer, R.A. and Greyser, S.A. Advertising in America: The consumer view. Boston: Division of Research, Graduate School of Business Administration, Harvard University, 1968.
- Becknell, J.C., Wilson, W.R., and Baird, J.C. The effect of frequency of presentation on the choice of nonsense syllables. Journal of Psychology, 1963, 56, 165-170.
- Burgess, T.D.G., and Sales, S.M. Attitudinal effects of "mere exposure": A reevaluation. Journal of Experimental Social Psychology, 1971, 7, 461-472.
- Heeler, R.M. A laboratory investigation of inter-related effects of mixed media, multiple copy, and multiple insertions in advertising campaigns. Unpublished doctoral dissertation, Stanford University, 1972.
- Light, M.L. An experimental study of the effects of repeated persuasive communications upon awareness and attitudes. Unpublished doctoral dissertation The Ohio State University, 1967.
- Miller, S.J., Mazis, M.B., and Wright, P.L. The influence of brand ambiguity on brand attitude development. Journal of Marketing Research, 1971, 8, 445-449.
- Olson, J.C. Cue properties of price. Working Paper No. 20, Pennsylvania State University, 1974.
- Orne, M.T. Demand characteristics and the concept of quasi-controls. In R. Rosenthal and R. L. Rosnow (Eds.), Artifact in behavioral research. New York: Academic Press, 1969, 143-179.
- Ray, M.L. and Sawyer, A.G. A laboratory technique for estimating the repetition function for advertising media models. Journal of Marketing Research, 1971, 8, 20-29.
- Rosen, N.A. Demand characteristics in a field experiment. Journal of Applied Psychology, 1970, 54, 163-168.

- Rosenthal, R. and Rosnow, R.L. (Eds.), Artifact in behavioral research. New York: Academic Press, 1969.
- Rosnow, R.L. and Aiken, L.S. Mediation of artifacts in behavioral research. Journal of Experimental Social Psychology, 1973, 9, 181-201.
- Rosnow, R.L., Goodstadt, B.E., Suls, J.M., and Gitter, A.G. More on the social psychology of the experiment: When compliance turns to self-defense. Journal of Personality and Social Psychology, 1973, 27, 337-343.
- Saegert, S., Swap, W., and Zajonc, R.B. Exposure context and interpersonal attraction. Journal of Personality and Social Psychology, 1973, 25, 234-242.
- Sawyer, A.G. Demand characteristics in laboratory experiments in consumer research. Working Paper No. 74-9, University of Massachusetts, 1974.
- Sawyer, A.G. The effects of repetition: Conclusions and suggestions about experimental laboratory research. In G.D. Hughes, M.L. Ray, and G. Haines (Eds.), Consumer information processing. Chapel Hill: University of North Carolina Press, in press.
- Stang, D.J. The effects of "mere exposure" on learning and affect. Journal of Personality and Social Psychology, 1974.
- Suedfeld, P., Epstein, Y.M., Buchanan, E., and Landon, P.B. Effects of set on the "effects of mere exposure." Journal of Personality and Social Psychology, 1971, 17, 121-123.
- Weber, S.J. and Cook, T.P. Subject effects in laboratory research: An examination of subject roles, demand characteristics, and valid inference. Psychological Bulletin, 1972, 77, 273-295.
- Zajonc, R.B. Attitudinal effects of mere exposure. Journal of Personality and Social Psychology, Monograph Supplement, 1968, 9, 2, Part 2, 1-27.
- Zajonc, R.B., and Rajecki, D.W. Exposure and affect: A field experiment. Psychonomic Science, 1969, 17, 216-217.
- Zajonc, R.B., Swap, W.C., Harrison, A.A., and Roberts, P. Limiting conditions of the exposure effect: Satiation and relativity. Journal of Personality and Social Psychology, 1971, 18, 384-391.





## VALIDITY PROCEDURES AT THE SURVEY RESEARCH CENTER

F. Thomas Juster  
Survey Research Center  
The University of Michigan

Besides the usual validity checks, the Survey Research Center uses tape recorded interviews as a means of observing interviewer-respondent interaction and to monitor interviewer performance. High levels of interviewer performance are also maintained by institutional arrangements: interviewers are regular University of Michigan employees and are compensated on a time and travel cost rather than a completed contract basis. Hence they tend to be relatively experienced and to have more on-the-job training than is usually true for survey organizations. SRC also maintains a highly professional sampling staff and uses rigorous probability sampling techniques.

Validity research at SRC includes work on ways to provide respondents with both incentives and cues. The former is designed to increase willingness to do the work required to provide accurate responses, the latter to create an interview situation maximizing the likelihood that the respondent will be able to recall accurately.

Discussions of validity procedures divide into two basic classes of cases: validity where an unambiguous external check is potentially available, and validity where no such external reference point exists. Most survey data fit into the first category, but at the Survey Research Center a very large proportion of survey measurements fall into the second.

An interview, to use the language of Kahn and Cannell in their classic, The Dynamics of Interviewing, is a specialized form of verbal interaction, initiated for a specific purpose, and focused on some specific content area with consequent elimination of extraneous material. Thus a valid response to a question on an interview simply represents one where the process of that interaction results in an answer which corresponds to the externally verifiable reality in the one case, and to an accurate description of the respondent's perceptions, opinions, or judgments in the second. Clearly however there is a major difference in strategy for these two cases: if one is trying to find out an objectively verifiable fact, whether the subject was a patient at a hospital or a clinic within the past year, whether he has a television set in the house, how many people live in the house, whether there is indoor plumbing, or how much income the family has--any strategy that will help produce truth is acceptable. But if the subject of the interview is whether the respondent thinks Ford will be a good president, or whether he expects prices to rise and if so, by how much, or whether he feels that he is fairly compensated on his job or that the level of public services in his neighborhood is satisfactory, then the exact specification of the question, the precise words and manner of the interviewer, will determine, to some extent, what kind of answer is forthcoming.

This paper is divided into three sections. In the first, I simply describe existing practice at the Survey Research Center. In the second, methodological studies that are underway at SRC are summarized briefly. In the third, I discuss the relation between validity and cost, and provide some generalizations (largely my own biases) about the payoff to validity checks and procedures and their role in the end product of survey research--the discovery and systematic accumulation of scientifically valid knowledge about behavior.

### Current SRC Procedures

Validation procedures at the Survey Research Center relate both to the instrument itself and to the performance of interviewers. For the instrument itself, we provide the usual package of routine validity checks--extensive pretesting, small scale pilot studies, debriefing sessions with the pretest and pilot study interviewers, etc. These can be viewed as essentially ways of insuring that the instrument is operational, in the sense that the questions as drafted are (or seem to be) clear to the respondent, and that the answers appear to be responsive to the intent of the question as seen by the researcher who designed it or helped to design it. The principal difference here between our procedures and others is probably that we do more of these tests and that the lag between initial pretest and final version is likely to be longer. That is principally because we have more relaxed time frames--or at least often do--than other organizations, particularly those that need to meet market demands for information on a schedule dictated by the demand side of that relationship.

Other procedures that are systematically incorporated into much SRC survey research relate to measures designed to observe trends and procedures designed to check for drift in the response to open-end questions. As indicated earlier, there are a great many questions and questionnaires in which no validity measure is really possible because what is being obtained is the perception, the opinion, or the judgment of the respondent. But even though there is no way to determine validity in an absolute sense, we often find it useful to repeat identical questions over different periods of time in order to obtain valid measures of differences.

Such differences are valid if and only if the stimulus presented by the question to the respondent is the same in fact, even though the question is identical. To validate that, we use random probes: Every  $n$ th respondent is asked a probe "Why is that?" question in order to determine whether the question is being answered with the same frame of reference as was true several years back. For example, questions concerning desegregation are quite likely to produce different responses now than 10 years ago, even if the respondents' attitudes are exactly the same, simply because the word itself tends to mean something different now than 10 years ago.

We have also conducted extensive experimental work designed to check instrument validity in situations where external reference points are available. These procedures and results are discussed below in the next section.

The last procedure I wish to note in discussing instrument validity is the

use of tape recordings as a way of observing interviewer-respondent interaction, problem questions, and so forth. Over the last several years, we have begun routinely to obtain tape recordings of pretest interviews. It is often true that one can judge problem questions--those where the meaning of the question to the respondent is obviously quite different from what was intended--by listening to the full interaction between the interviewer and the respondent. Tape recordings are also used, as discussed below, as a means of monitoring interviewer performance.

### Interviewer Quality Control

Possibly the most important single dimension of validity procedures in survey operations is to insure that operational data obtained from the field have the same characteristics as the best data that can be obtained from the survey instrument by a skilled, highly motivated, and well trained interviewer. This does not really bear on the issue of whether the instrument is able to obtain valid data--it simply insures that, whatever the potential validity of the instrument, the actual validity of the results comes as close as possible to that.

In this area, the Survey Research Center uses a wide collection of techniques, many of which are standard and found uniformly elsewhere and some of which are probably not. Among the more or less standard ones are:

- (1) Skilled interviewers routinely conduct reinterviews with respondents, and the results are compared against the original interview. Persistent flaws in interviewer performance are thus monitored.
- (2) Careful attention is paid to quality control throughout the entire survey process--interviewer selection and training, the preparation of materials for the interview, substantive discussions with interviewers by the study directors, pre-study conferences after training interviewers, and sampling procedures are all given a good bit of attention. Most of what is involved here is based on interaction between the interviewers and the research staff, the research staff and the sampling staff, etc.
- (3) Many of our studies involve a reinterview of sample households, and all of our interviewers are aware of this fact.
- (4) A sample of respondents is routinely mailed a follow-up questionnaire, primarily to check on the time a survey schedule takes to complete. Interviewer shortcutting may take the form of mixing a small amount of real information with fictitious data, with the result that total time will be substantially less for partly fictitious interviews.
- (5) Computerized data on costs per interview, average time per interview, etc. are routinely made available to the field staff, and interviewers whose costs are at one or the other range of these distributions are thus identified regularly. Poor interviewing technique is apt to be associated with either very

high cost or very low cost relative to the mean.

There are a number of aspects of SRC field operations which are to some degree unique to the organization, and which bear on the quality of interviewer performance and thus on the validity of data.

These procedures in large part have to do with interviewer selection and motivation, rather than the specifics of interviewer training on any given study. Others would be generally applicable to any survey organization.

The first point to note is that SRC interviewers are classified as regular University of Michigan employees, rather than as independent contractors. This means that interviewers have Social Security taxes withheld, and it also means that interviewing costs are subject to University overhead. The latter has both good and bad aspects, since it raises budgeted costs as well as providing overhead funds for general use in the Center. Largely because of the role of SRC as part of the University, we may also gain some advantage in terms of interviewer motivation with respect to training, work performance, etc. For the most part, the subject matter content of SRC studies deals with what most people would construe as important social issues, where the level of interviewer interest might well be higher than in, for example, studies of marketing habits or brand preferences. In addition to their status as regular U of M employees, SRC interviewers are typically paid on a time and travel cost basis, rather than a completed contract basis. That imposes costs, or at least would have that tendency, but it also has, in our view, significant payoffs in terms of the quality of completed interviews.

As a consequence of these contractual arrangements, SRC interviewers have a number of characteristics which are measurable and which can be presumed to contribute to the validity of the survey measures:

- (1) The staff is relatively stable--average experience with SRC is 5 years.
- (2) Because we have the expectation of greater staff stability, it pays us to invest more than otherwise in on-the-job training for interviewers. It also means that we have an opportunity to retain interviewers who are at the top end of the performance range.
- (3) Both staff stability and training means that we have less need for extensive editing of completed interviews, once a check for general acceptability has been made. That is more of a cost saving than a quality control measure, although it is made possible by the experience and stability controls.
- (4) Finally, the pay-for-time and cost system provides no disincentive to obtain interviews that are difficult to obtain. The result should be a more representative sample of the universe as it actually exists, with consequent impact on the validity of the data.

Two other characteristics of our general field operation should be noted.

As indicated earlier, we have been making extensive use of tape recorded interviews as a way of pretesting questions, and these techniques are now being used as a routine check in the evaluation of interviewer performance. Before any interviewer can be raised to a higher pay level, a supervisor report based on tape recorded interviews is mandatory. Our experience here is quite interesting. For example, every survey organization develops folklore about which interviewers "do a good job" and which interviewers one might have doubts about or who appear to be, by some subjective standard, sub-par in their performance. It often turns out, however, that the folklore is wrong. Interviewers who have a great deal of experience and a relatively high standing among peers are sometimes found to be interviewers who "take over" an interview situation and produce a set of results which is as much a consequence of the interviewer's actions as it is the respondent's perceptions. Listening to a tape recording of the completed interview enables our supervisory personnel to measure such straightforward aspects of data quality and validity as how the questions are asked (whether the way they were written or some nonlinear transformation thereof), whether interviewers use directive probes or nondirective ones, whether skip sequences are handled appropriately, whether the pace of interviewing is such as to be conducive to accurate responses, etc. On some of these issues, we have also done some experimental methodological work--for example, on the question of how the "pace" of the interview relates to response validity. But the principal point is that periodically listening to the tape recording of an actual interview by one of the field interviewers is an exceptionally good way to nip bad habits in the bud, to be sure that interviewers are doing what they are supposed to be doing and not something different, and in general to control on quality so as to insure that the data are as close as they can be to what the questionnaire instrument is capable of providing.

Like everyone else, I also have my own favorite tales of what happens to data quality when control over interviewers is casual or, through force of circumstances, suboptimal. One recent case in point, or at least my interpretation of it, suggests the dimensions of the problem.

For a number of years, the U.S. Bureau of Census has been conducting household interviews concerned with consumer purchase plans for automobiles, houses and major durables. Around the middle of the 1960's, some research done by myself and others suggested that a more valid measurement of purchase expectations might well be obtained from a questionnaire which focused on subjective probabilities rather than on something called plans or intentions. We did extensive pretesting of this general concept, including one final pretest in which a random sample of households was interviewed with both the plan or intention version and the subjective probability version, the interviews being conducted a few days apart so that the "real" value of the anticipatory variables was presumably unchanged. Results from the pretest were clearcut: in explaining differences among households in purchase behavior, subjective probabilities were overwhelmingly superior to purchase plans or intentions. Put into a multivariate equation, purchase plans washed out entirely, leaving subjective probabilities the dominant variable.

A major deficiency of the purchase plan approach--the tendency for most actual purchases to be made by households who did not report plan or intention--was substantially ameliorated by the subjective probability approach, in that the portion of purchases made by households with non-zero probabilities

was significantly larger than it had been for the plan or intention version. The mean values of the subjective probability scale, which if taken literally should be a direct forecast of the actual purchase rate, were quite close to observed purchase rates for the sample households. Thus on all counts that we could think of, the evidence was convincingly clear that subjective probabilities were a better way to predict purchase behavior.

Hence the Census survey was changed in 1967, with subjective probabilities replacing the purchase intention variables. However, subsequent experience with the use of these data to predict changes in purchases over time do not indicate any superiority at all for the purchase probability measure, and if anything, indicate that it may well be inferior to the plan or intention variable. And in fact, the Census Bureau recently (July 1973) decided to discontinue the buying expectations survey because they felt (whether rightly or wrongly is not the point) that the survey made an insufficient contribution to purchase forecasts to warrant its relatively high cost.

Aside from the inability of the probability measure to improve on the predictive value of the plan or intention measure in time-series analysis, an interesting aspect of the subsequent development is that the structure of the probability data in the operating survey never really approached the structure found in the initial pretest. In particular, it tended to be true that the proportion of total purchases accounted for by non-zero probability households in the operating survey was quite close to the proportion previously found for planners or intenders, rather than substantially larger as the pretest had suggested it would be and as theory suggested it ought to be.

There are many possible explanations for this set of results. One, which is not really possible to explore in retrospect, is that interviewer treatment of this survey instrument essentially invalidated its usefulness. The Census Bureau, as all of you know, concentrates largely on the measurement of objectively verifiable phenomena--whether or not a person is unemployed, whether there is indoor plumbing in the house, whether the household owns a car, whether the house is owned or rented, etc. Census has relatively limited experience with the so-called soft or subjective measures--in general, with the broad area of perceptions, opinions, expectations, etc. The pre-study conferences, some of which I attended, made it quite clear that many Census interviewers regarded asking households about the subjective probability of their making a purchase as an affront to the intelligence of both the interviewer and the respondent. My own experience with this kind of measure is that the interviewer has to be extremely careful not to put an answer into the mouth of the respondent, since the question is of course quite a difficult one and tends to suggest a degree of precision that interviewers and respondents alike are apt to think is unrealistic. But that is not the point, of course, since uncertainty about a probability judgment by no means renders those judgments invalid when aggregated across households. But skepticism on the part of the interviewer about the usefulness of the survey instrument surely can render these judgments invalid, and the inability of the probability data to predict actual purchase behavior, as well as the marked difference in structure between the operating survey and the pretest, suggests that some of this skepticism may have rubbed off on the survey itself and hence on the validity of the data.

### Sampling Procedures

An important dimension of survey research which bears strongly on predictive validity--as distinct from the narrower question of the validity of a particular response from a particular respondent--has to do with the procedures by which households are selected and the additional procedures by which designated respondents are the ones to be interviewed. The Survey Research Center probably spends more time fussing about sample selection and obtaining interviews with designated respondents than any other organization outside the Census Bureau. All of our samples are multi-stage probability samples with very limited clustering. Interviews are conducted only with designated respondents, and no substitutions are allowed. One consequence of these comparatively rigid techniques is that we tend to have relatively high nonresponse and, at the same time, relatively high costs.

I had an interesting experience with just how well the sampling operation is done at SRC, in connection with a recent project on the impact of the federal government's General Revenue Sharing Program on state and local governments. To conduct the study, we had to select a sample of U.S. municipalities. Such samples apparently do not exist--except for what the Census Bureau euphemistically calls a "sample" consisting of some 15,000 of the 38,000 total U.S. municipalities. We started with our sampling frame of PSU's, which is of course designed to represent people and not governments, and proceeded to draw a stratified sample of municipalities within some seven city size classes. After any number of modifications of the ongoing sampling frame (new counties had to be added in areas where the municipality sample was especially thin), the final sample was eventually specified and assigned the appropriate weight for its city size class, probability of selection within the PSU and the probability of selection for the PSU as a whole.

The adequacy of this sampling frame can be judged quite accurately, since we have a tape prepared by the Office of Revenue Sharing which includes every U.S. county and municipality and provides data on tax collections, population, and Revenue Sharing allocation. Comparisons against the universe suggests that this 800 municipality sample provides a remarkably close representation of all U.S. counties and municipalities, and of both regional and size class distributions. The error in generating aggregate statistics from this sample, even for regional breakdowns, is apparently going to be no more than a few percent. Although having a carefully drawn probability sample will clearly do nothing to improve the validity of individual responses, it goes a long way to insure that validity at the individual response level can be translated to accurate descriptions of the population as a whole.

### Validity Research

Some years back the Institute for Social Research, the "holding company" that includes the Survey Research Center as well as three other quasi-independent groups, instituted a research program designed to find ways of improving the quality of information reported in personal interview surveys. Over the years there has been substantial technical progress in most phases of survey research--sampling, statistics, methods of analysis, use of computers, etc. As noted in a recent working paper by Charles Cannell, who heads up this



program at the Institute, "The 1973 model of the survey interview does not vary to any significant extent from the 1943 version."

In part, the reason for the technological backwardness of interview methodology in the social sciences is that most of us work with social-psychological variables which lack an objective external reality against which the validity of survey responses can be compared and evaluated. One can of course be tautological--a valid attitude, perception or expectation is that which is reported to an interviewer on a survey. But if this were really so, characteristics of the interviewer--such as age, sex, race, socio-economic status, attitudes and preconceptions--would be uncorrelated with the reporting of social-psychological variables. We know that this is not the case, though, and hence the problem cannot be assumed away in this manner.

Much of the current program of research on validity at SRC is concerned with what is clearly the simpler of the validity issues--how to obtain accurate reporting where we can test experimental procedures against an objectively verifiable reality. The experimental work has been done almost entirely in the health area, thanks to a series of contracts and later grants with the National Center for Health Statistics and the National Center for Health Services Research and Development.

In these studies, the major dependent variable is some health attitude or behavior--chronic and acute illness, injury, medication, use of health facilities, reaction toward medical care, etc. Note that data in the health field have the twin characteristics that (a) they are often capable of being verified from medical records, and (b) even when overtly factual, they contain a good deal of latent emotional and attitudinal content. The remainder of this section summarizes and highlights the results obtained in these methodological studies.

#### Memory Failure

The model of an interview situation with which Cannell and his coworkers at SRC operate is that an interview situation is one in which interaction between interviewer and respondent has the purpose of stripping away nonessentials (for the interview) and exposing information possessed by the respondent that is relevant to the purposes of the interview. Thus one issue is whether or not the respondent can be programmed to improve recollection and reporting of specific events. In general, the problem in health surveys typically shows up as underreporting, although in principle overreporting would be possible if biases other than memory failure are introduced. The research provides wholly unambiguous results, which are consistent with the findings of others: the longer the duration between the interview and the time of the event (hospitalization in this case), the less likely the respondent to report the event (having been hospitalized). This is not surprising, and as noted is probably one of the few firmly documented methodological conclusions in validity studies. What is surprising is the rapidity with which the reporting decay reaches major proportions: underreporting of actual hospital or clinic visits gets to be almost half the total of known visits after about a year. (Table 1)

A comparable dimension to duration is salience--events which are noteworthy in the total stream of the respondent's experiences tended to be reported much more often than other events. The duration of hospital visits

TABLE 1

## Underreporting and the Dating of Events

A. Hospitalization Reports by Number of Weeks between Discharge and Interview  
(includes proxy respondents)

<u>Number of Weeks between Hospitalization &amp; Interview</u>	<u>Number of Hospitalizations (From Hospital Records)</u>	<u>% Not Reported in Household Interview</u>
1-10	114	3
11-20	426	6
21-30	459	9
31-40	339	11
41-50	364	16
51-53	131	42

B. Chronic Conditions Reports by Number of Days since Last Visit to Clinic  
(SRI)

<u>Number of Days since List Clinic Visit</u>	<u>Number of Conditions in Records</u>	<u>% Not Reported in Household Interview</u>
1-7	116	9
8-14	218	28
15-28	440	24
29-56	683	42
57-84	574	37
85-112	513	42
113-140	476	45
141-168	355	46
169-224	372	57
225-280	1,232	52
281-364	1,078	58
365+	71	59

SOURCES: "Reporting of Hospitalization in the Health Interview Survey," Cannell, Fisher, Bakker. Vital & Health Statistics, U.S. Public Health Washington, D.C., July 1965, Series 2, No. 6, and "Interview Data on Chronic Conditions Compared with Information Derived from Medical Records," Madow, Vital and Health Statistics, U.S. Public Health Service, Washington, D.C., May 1967, Series 2, No. 23.

or the frequency of clinic visits are the measure of salience, and the relationships between fraction of events reported and salience are shown in Table 2. About 25 percent of one day hospitalizations were not reported in the interview, and over half of single visits to clinics failed to be reported. (These two kinds of memory issues give attitude results.)

### Refusal Rather than Forgetting

A second dimension of invalid responses on interviews concerns the respondent's unwillingness to provide information, rather than inability to do so. Here, the experimental results concern the probability that reporting a particular diagnosis will be a source of embarrassment, or more directly, responses to the question: "If you had x, how willing would you be to have other people know about it?" Cannell finds that a direct measure of embarrassment--percent willing to report if they had x--corresponds remarkably closely to the fraction of valid reports for those same conditions generated by externally validated and independent data. (Table 3)

What these results suggest is a model in which the interviewer-respondent interaction comes closest to providing valid data when, first, the event on which survey data are to be obtained is one that the respondent has in the forefront of his total recollection of events, and second, the event to be reported is not regarded as threatening to the respondent's self-esteem. Thus the question arises: can one influence the behavior of the respondent by manipulating the behavior of the interviewer in such a way as to change the characteristics of the interaction so as to enhance validity?

### Interviewer Respondent Interaction

Since the interview is modelled as an interactive situation, how to change respondent behavior by manipulating interviewer behavior becomes the key issue. Cannell and his colleagues consequently taped a number of household interviews, and had the interviews categorized into respondent and interviewer behavior. For example, does the interviewer read the question as written, does he use nondirective probes, does he capture the respondent's answer when he notes the response, etc. For the respondent, is clarification requested, is the answer appropriate to the question, is there refusal, etc.

The analysis indicated a high degree of positive correlation between the total number of units of activity of both the interviewer and the respondent. One would have hoped that reluctant respondents would be cajoled out of reluctance by active interviewers, but this turned out not to be the case. The data also indicated that interviews where both respondents and interviewers showed a high level of "positive" task behaviors tended also to show low levels of nontask relevant behaviors. Thus one possibility is that the respondent follows the pattern set by the interviewer, and if one can change interviewer behavior one can manipulate respondent behavior.

Of particular importance in the interaction process is the nature of feedback used by interviewers. The data indicated that about a quarter of total activity on the part of interviewers constituted feedback--short interjection, longer comments (That's the kind of information we want...), etc. But the data also clearly showed that feedback was not related to the adequacy or inadequacy of response. Instead, the customary use of feedback seemed to

TABLE 2

## Underreporting and the Intensity of Events

## A. Reports of Hospitalizations by Duration of Hospitalization (including proxy respondents)

<u>Duration of Hospitalization (in days)</u>	<u>Number of Hospital Records</u>	<u>Percent Not Reported in Household Interview</u>
1	150	26
2-4	646	14
5-7	456	10
8-14	352	10
15-21	111	6
22-30	58	2
31 and over	46	8

## B. Reports of Chronic Conditions by Number of Visits to Clinic (SRI)

<u>Number of Visits to Clinic</u>	<u>Number of Conditions</u>	<u>Percent Not Reported</u>
1	3,081	56
2	1,281	47
3	643	35
4-5	639	26
6 or more	496	14

SOURCES: "Reporting of Hospitalization in the Health Interview Survey," Cannell, Fisher, Bakker, Vital and Health Statistics, U.S. Public Health Service, Washington, D.C., July 1965, Series 2, No. 6, and "Interview Data on Chronic Conditions Compared with Information Derived from Medical Records," Madow, Vital and Health Statistics, U.S. Public Health Service, Washington, D.C., May 1967, Series 2, No. 23.

TABLE 3

Subjective Willingness to Report Medical  
Conditions and Percentage of Cases Actually Reported

<u>Conditions</u>	<u>% Willing to Report (79 students)</u>	<u>% Valid Reports in Household Survey (HIP)</u>
<b>More serious conditions:</b>		
1. Asthma	84	71
2. Heart disease	58	60
3. Hernia	55	54
4. Malignant neoplasm	31	33
5. Mental disease	19	25
6. Genito-urinary disease	14	22
<b>Less serious conditions:</b>		
1. Sinusitis	89	48
2. Indigestion	88	41
3. Hypertension	83	46
4. Varicose veins	65	42
5. Hemorrhoids	21	38

**SOURCE:**

"Some Thoughts about Interview Data," by Sidney Cobb and Charles F. Cannell,  
International Epidemiological Association Bulletin, 13, 1966, 43-54.

be comments made after refusals, presumably as means of building or restoring rapport. It turned out that feedback was not in general used to "reward" positive responses, nor was the absence of feedback used to "penalize" inadequate responses.

The findings from the interaction analysis led to a series of experiments designed to control and use feedback as a positive force to improve reporting, and to capitalize on those results showing a balance in the level of interactions during the interview, with the level being highly correlated with the amount of information reported. The model specifies that: (1) positive feedback will tend to produce better information, and (2) a high level of verbal activity by the interviewer will tend to produce an equally high level on the part of the respondent, with more and better information resulting from these higher levels of verbal activity.

The feedback experiment was conducted by instructing interviewers to provide a reinforcing statement ("That's the kind of information we need," "That's useful information," etc.) after each positive report by the respondent, with several such statements being prepared and the interviewer using them in sequence. The verbal activity experiment was conducted by designing long and short versions of essentially identical questions, for example:

"Have you ever had any trouble hearing?" (short form)

"Trouble hearing is the last item on this list. We are looking for some information about it. Have you ever had any trouble hearing?" (long form)

The results of these experiments were as predicted--reinforcement produced greater validity, as did longer questions. However, when the two procedures were combined with the expectation that they would be reinforcing, both showed main effects, but the combination of the two showed lower reporting rates than either technique by itself. Analysis of these data by respondent educational level suggests the answer, as indicated in Table 4. In a nutshell, reinforcement improves reporting for less well educated respondents, but not for more highly educated ones. But for respondents with less education, short questions produced better results than long ones, and vice versa for highly educated respondents. Overall, for respondents with less education short questions with reinforcement works best, while for more educated respondents, long questions without reinforcement is optimal. One can easily find plausible interpretations of these findings. Less well educated respondents tend to rely more than others on interviewer cues to direct reporting. Thus feedback and reinforcement aids performance for this group, and since the interviewer is apt to be higher status than the respondent, feedback is even more appropriate and welcome. But more highly educated respondents do not need reinforcement and even perceive it as inappropriate and condescending.

On the other hand, long questions, simply because they are long, tend to confuse less well educated respondents, but highly educated respondents tend to benefit from expressing themselves more fully to reflect nuances of thought, from interviewers' high verbal activity, or from the clearer understanding of the specific tasks to be performed that is permitted either by the language or by the additional thinking time given them by the longer questions.

TABLE 4

Percent of Chronic Conditions Identified by Physicians  
Reported in the Interview:  
By Respondent Education and Experimental Treatment

<u>Interview Procedure</u>	<u>Education of Respondent</u>	
	<u>Less Than High School Graduate</u>	<u>High School Graduate and Above</u>
Short Questions:		
Without Reinforcement	61%	56%
With Reinforcement	77	42
Long Questions:		
Without Reinforcement	56	69
With reinforcement	65	54

There are between 36 and 53 persons per group.

SOURCE:

"Reporting Health Events in Household Interviews: Effects of Reinforcement, Question Length, and Reinterviews," Marquis, Cannell and Laurent, Vital and Health Statistics, U.S. Public Health Service, Washington, D.C., March 1972, Series 2, No. 45.

### Some Final Thoughts on Validity

For the most part, the discussion above has been in terms of validating survey measures that are relatively simple to measure and extract from the respondent's memory bank. Motivation and willingness to exert effort in gathering the information is probably a more important source of difficulty than inability to recall. But in at least one major area of survey research, the validity problem is at least as much a memory as a motivation problem--the collection of data from households on financial flows and asset-debt holdings.

One interesting direction that needs to be pursued here is an increased reliance on objective data that the respondent can obtain from records rather than an interview situation where recall is the only possible basis for information. Several issues are involved--do the data sources exist, can the respondent be persuaded to use them, are they sufficiently comprehensive so that they can form the basic raw materials for survey responses, can they be judiciously blended with recollection to provide a richer and more accurate data set, etc.

We have had enough experience over the last several decades to be able to form judgments about the validity of survey data on consumer financial flows and asset-debt holdings. As a generalization, accurate financial flows cannot be recalled, even with the best will in the world and with no concern over disclosure. This is less true for families whose financial affairs are very straightforward--they have no assets but a savings account, they have the same income flow every week or month, they have a spending pattern that usually exhausts all of their receipts and goes for a repetitive set of items, etc. But that kind of financial flow pattern is less and less common as the society becomes richer, more people enter the labor force working part time hours, more people have assets in different forms and several kinds of debts, etc.

It seems clear enough that the solution to these kinds of problems is to obtain access to the financial records possessed by households. And here the principal difficulty is likely to be willingness rather than ability, since one of the features of any increasingly rich society is the greater profusion of paper that seems essential to the way we function. Only a few (about 15%) of U.S. families still lack checking accounts; only a slightly higher percent do not have savings accounts. The vast majority of income payments are received in the form of checks which must be cashed or deposited; interest, dividend, and many nonregular forms of income are, as a matter of statute, now reported to the household every year by way of information forms which are designed to be used as a basis for income tax filing; credit card companies regularly send monthly statements of charges and payments, as well as interest costs; and so forth.

On the basis of these kinds of financial records, it appears to be true that a comprehensive survey of household financial status (income flows, expenditure flows, asset and debt holdings) could be conducted largely on the basis of records which presumably form a much more precise source of data than any alternative.



A second kind of issue bearing on validity in the broader sense is that of costs. It seems unproductive to think of validity only in the narrow sense of the relation between the true value of the desired measure and the value obtained from a survey. Validity has a price, and the question must be faced: what is the optimum measure of validity in a cost-effective sense? In this context, by cost effective I simply mean the degree of validity which just pays its own way in terms of reducing the uncertainty that would attach to measures of lesser validity by enough to justify the additional costs. In economists' jargon, the marginal revenue from an increment of validity must be at least as high as the marginal cost of obtaining that increment in order to justify, either for private or social purposes, the use of resources to increase validity to any given point. Unfortunately for neat or simple analysis, there are considerable opportunities for tradeoffs as regards validity in this very broad sense. For example, the validity of a psychological scaling measure is dependent on the ability of that measure to classify households into appropriate subgroups--those with the same response within group, but different responses from other groups. Validity could be increased by measures ranging from investing heavily in a methodological research design to improve our understanding of what the scale reflects to simply adding additional scales that capture similar dimensions of perceptions or attitudes and creating groups with a cluster of such scales. To really push validity to its optimum point--to decide what it's worth to improve our classes--we have to know how much gain would be achieved by each of the possible methods and how much cost is involved, then determine the most efficient way to proceed, then decide how much we really gain given the purposes for which the scale is to be used. Although I do not have much that is concrete to offer in terms of criteria, it does seem to me important to recognize that higher validity carries a price tag and may represent a misallocation of research resources.

## VALIDITY AND GOODNESS OF FIT IN DATA ANALYSIS

Donald R. Lehmann<sup>1</sup>  
Columbia University

This paper describes the relationship between the truthfulness and usefulness of a model or hypothesized relationship between two or more variables and the goodness of fit which results from an empirical investigation of the model or hypothesis. Several obvious discrepancies between model "validity" and goodness of fit results due to such causes as measurement problems, model mis-specification, and stochastic behavior are described. In addition, a procedure for assessing the appropriateness of a linear regression model is presented.

### Introduction

The first step in formal research is often the specification of a model, either explicitly or implicitly. Data is then collected and the model rejected if the data and the model conflict. In the social sciences, this type of model investigation has been labeled predictive testing (Basman, 1965). This approach consists of deducing conclusions from the model and seeing if the data supports the conclusions.

The predictive testing approach, in spite of its obvious appeal, is not used widely in most research on consumer behavior. One reason for this is that in much research, a model is not extant prior to data analysis. Even when a model is present, it is often poorly specified. For example, A could be hypothesized to:

- 1) be related to B
- 2) be related to B in some specific mathematical way (eg.,  $A = c + dB$ )
- 3) be related to B in some specific mathematical way with constraints on the parameters (eg.,  $A = c + dB$  where  $c > 1$  and  $0 < d < 1$ )

Moreover, it is possible to specify the relationship between A and B to be causal or merely correlational and to assume a large random component exists. Hence lack of certainty over what the model is makes predictive testing very difficult.

Even when a model is specified, the predictive testing approach is often difficult to utilize. In the first place, it requires considerable ingenuity to deduce meaningful conclusions from models which can be falsified by data. This is especially true for models which do not assume causality. Secondly, in many cases several competing models exist, and with notable exceptions (Bass & Clarke, 1972), more than one of them can fail to be rejected. Hence because of difficulties with the predictive testing approach, as well as different training, most research on consumer behavior has tended to focus on goodness of fit measures as an indicator of useful relationships.

Goodness of fit measures have some obvious advantages in model development. First, they are quantitative indices which can be compared across models so that the best of a set of models can be selected. Secondly, there exist established statistical procedures for testing these measures for significance. Third, they can be used as a means of deducing new or modified models from a set of data. Fourth, they appear as output in canned computer programs, which is no small reason for their prominence in reported research results. Yet in spite of their advantages, goodness of fit measures have some important limitations as indicators of the truthfulness and usefulness of models.

#### Limitations of Goodness of Fit Measures

It is very easy to rely on goodness of fit measures as a means of evaluating model usefulness. Unfortunately, there are a variety of reasons why perfect goodness of fit measures should not be expected from useful models. These reasons include:

##### Imperfect Model Operationalization

One obvious situation where goodness of fit measures are not appropriate for estimating a model's usefulness is when the operationalization of the model does not match the model. In other words, the operationalization of the theoretical model may be due to the type of data available for examining the model. Criticisms of research on this basis are widespread (Cohen, et al., 1972; Taylor & Gutman, 1973). On the other hand, a high goodness of fit measure suggests that a useful model which is equally or even more valuable than the original model may be developed because of its empirical goodness of fit.

##### Sample Limitations

Goodness of fit measures' ability to estimate a model's usefulness are obviously limited to the population represented by the sample. Hence, for example, if for some reason college students fit a model differently from the general population, studies of "40 college sophomores" are potentially misleading.

##### Measurement/Scale Problems

A variety of measurement problems occur which will affect the size of goodness of fit measures. For example, individuals may not treat an apparently equal-interval scale as such (Stevens, 1946). In many other situations, a lower order scale (eg., ordinal) will be treated by the researcher as a higher order scale (eg., interval or ratio). In situations where non-interval scales are used in regressions as though they were interval scales, the practical maximum for  $R^2$  even assuming a valid model is often far less than 1 (Lehmann, 1972; Morrison, 1973, 1972a, 1972b). On the other hand, questionnaires may induce consistent responses across questions. When such consistency bias is present, the goodness of fit measures among the affected variables are artificially increased.

##### Minor Model Mis-Specification

It is possible that a model be correctly specified in the sense of

containing the correct constructs and links among constructs while incorrectly specifying the mathematical form of the relationships. In such cases, reliance on goodness of fit measures may lead to inappropriate conclusions.

For example, assume attitude is determined by a multi-attribute attitude model which is multiplicative with unequal weights on the attributes:  $A_j = \prod_{i=1}^n |B_{ji}|^{W_i}$  where  $B_{ji}$  represents the distance from the desired amount of the  $i^{\text{th}}$  attribute possessed by alternative  $j$ ,  $W_i$  is the importance weight of the  $i^{\text{th}}$  attribute, and  $A_j$  is the individual's attitude toward the  $j^{\text{th}}$  alternative. In this case, application of the more popular linear model ( $A_j = \sum_{i=1}^n |B_{ji}|^{W_i}$ ) can lead to the conclusion that equal weights are superior. Assume two alternatives, M, which is one unit from the desired level on attribute 1 and four units from the desired level on attribute 2, and N, which is 4 and 2 units from the desired levels on the two attributes. According to the true model (multiplicative with unequal weights), M is the preferred alternative. If the linear model is applied, however, the assumption of the true weights suggests that M and N are equally preferred while the use of equal weights correctly identifies M as the preferred alternative. Hence reliance on goodness of fit would falsely suggest that the importance weights for the two attributes were equal.

		<u>Additive Model</u>				<u>True Model</u>	
		Equal Weights		True Weights		True Weights	
		$(W_1=W_2=1)$		$(W_1=2, W_2=3)$		$(W_1=2, W_2=3)$	
<u>Alternative</u>		<u>Attitude Rank</u>		<u>Attitude Rank</u>		<u>Attitude Rank</u>	
M	(1, 4)	5	1	14	tie	64	1
N	(4, 2)	6	2	14	tie	128	2

A similar problem exists with comparing linear and non-linear forms of relationships. Given the degree of error or randomness in most survey research data (Hulbert & Lehmann, forthcoming), the property of a Taylor series guarantees that most of the relationship between two variables will be accounted for by the first (linear) term. Hence, unless orthogonal polynomials are employed, second and higher order terms will not appear to be very significant due to their high collinearity with the linear term. For this reason, goodness of fit measures tend to indicate that relations among constructs are essentially linear when in fact the relations may not be.

### Spurious Correlation

It is very possible that a model will produce a high goodness of fit while being essentially unrelated causally. This can happen because of a spurious correlation among two variables (Blalock, 1964). Hence, two variables A and B can appear related since a third variable C affects both A and B. Unfortunately, the goodness of fit level gives no indication of whether or not a spurious relationship exists.

### Inappropriate Application

In many cases a potentially valid model may be applied in an inappropriate way. For example, estimation of multi-attribute models across individuals by means of regression, though popular (Wilkie & Pessemier, 1973), seems an inappropriate application of an individual model (Beckwith & Lehmann, 1973). In such cases, the goodness of fit of the model does little to indicate its usefulness.

### Post-Hoc Models

In many situations, a model is essentially developed from data rather than tested by it. Hence the goodness of fit of such a model is biased upward. This is especially true when stepwise procedures are employed.

### Degree of Freedom Limitations

Goodness of fit measures are obviously biased upward whenever many parameters are estimated in relation to the size of the sample. The obvious remedies to this problem include the reporting of adjusted  $R^2$ 's and using split-half procedures (Morrison, 1969). Unfortunately, split-half procedures require a relatively large sample size which if extant tends to make the goodness of fit bias fairly small.

### Stochastic Influences

Many individuals have recently focused attention on the proposition that a considerable portion of individual behavior is random and hence in principle not explainable (Bass, 1974). This means that models with low goodness of fit indices may be both true and useful (Bass, et al., 1968). It also suggests that models with extremely high  $R^2$ 's may be modeling error or random behavior rather than the true underlying process.

### The Appropriateness of $R^2$

Probably no measure of goodness of fit receives more attention than the  $R^2$  in regression analysis. Most people feel that a good  $R^2$  is a big  $R^2$ , largely due to the influence of economists who are used to dealing with aggregate time-series data. This section will proceed to argue that  $R^2$  is a poor indicator of whether a linear model is useful and appropriate.

Much of the research on consumer behavior has resulted in  $R^2$ 's in the .05 and .10 range. As such, it had indicated little individual-level predictive power but significant relationships among variables such as income and TV viewing time. These variables are usually measured on a 5-8 point scale. Hence it is not at all surprising that people with incomes between 5 and 10 thousand dollars vary considerably in the amount of time they spend watching TV. Figure 1 represents a typical situation in survey research.

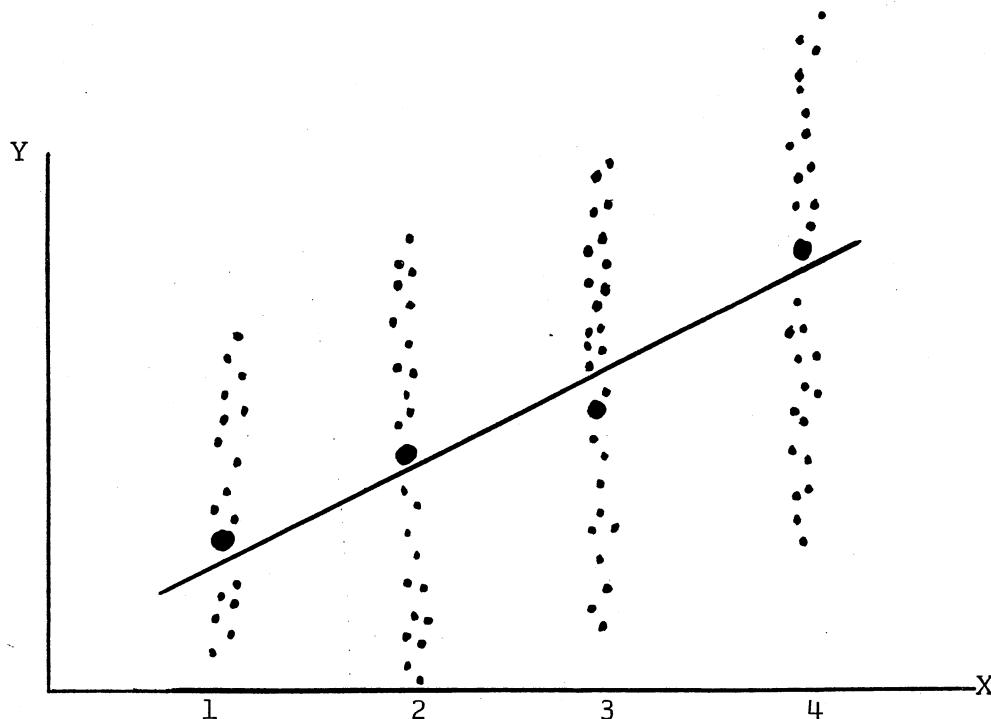


Figure 1. Hypothetical regression example.

At this point the question of exactly what model is being tested becomes important. If the model is that income is a deterministic predictor of TV viewing, then the model can be rejected. If, on the other hand, the model says TV viewing is related to income with a large random component, then the  $R^2$  becomes a poor indicator of the model's usefulness (Bass, et al., 1968).

Assuming the model is stochastic, the real question becomes whether average TV viewing differs across several income groups. This is essentially an analysis of variance problem. An interesting related question, however, is whether income and average TV viewing are linearly related. In order to address this question, it is useful to first partition the variance in the normal manner:

$$\text{Total variance} = \text{Within variance} + \text{Between variance}.$$

Next, by partitioning the Between variance we achieve the following situation (Lehmann, 1974):

$$\text{Total variance} = \text{Within variance} + \text{Regression Explained variance} + \text{Regression Unexplained variance}.$$

Recalling Figure 1, the within variance (which is generally by far the largest component of the variance) is essentially random and unexplainable by the independent variable. Since  $R^2$  is:

$\frac{\text{Regression Explained Variance}}{\text{Total Variance}}$ , it is necessarily low. A more suitable measure of the appropriateness of a linear model would be:

$PR^2 = \frac{\text{Regression Explained Variance}}{\text{Between Variance}}$ . This  $PR^2$  would be 1 if the means of the segments fell on a straight line, and hence it measures how well a linear function estimates the effect of the independent variable on the dependent variable.

Two other points about  $PR^2$  are important to recognize. First,  $PR^2$  is the  $R^2$  which would result if the mean behavior for each segment were regressed against the value(s) for each segment on the independent variable(s) with the means weighted by the number of observations in each segment. Second, departures from 1 by  $R^2$  can be tested to see if a significant non-linearity exists in the relationship. The null hypothesis of a linear relationship is tested as follows:  $F = \frac{\text{Regression Unexplained Variance}/df_1}{\text{Within Variance}/df_2}$ . Derivations for these results are available (Lehmann, 1974).

#### An Empirical Example of $PR^2$

This example is based on a mail survey of 513 homemakers taken in 1968. The relationship between average TV viewing (measured on a 7-point scale) and income (measured on a 6-point scale) appears as Figure 2. From this figure, it appears that income and TV viewing are linearly related. A simple regression produces an  $R^2$  of only .048, however, which is neither very encouraging nor enlightening. In order to see if the relationship is really linear,  $PR^2$  was calculated (Table 1). The results indicate a  $PR^2$  of .878, which is close to 1. The

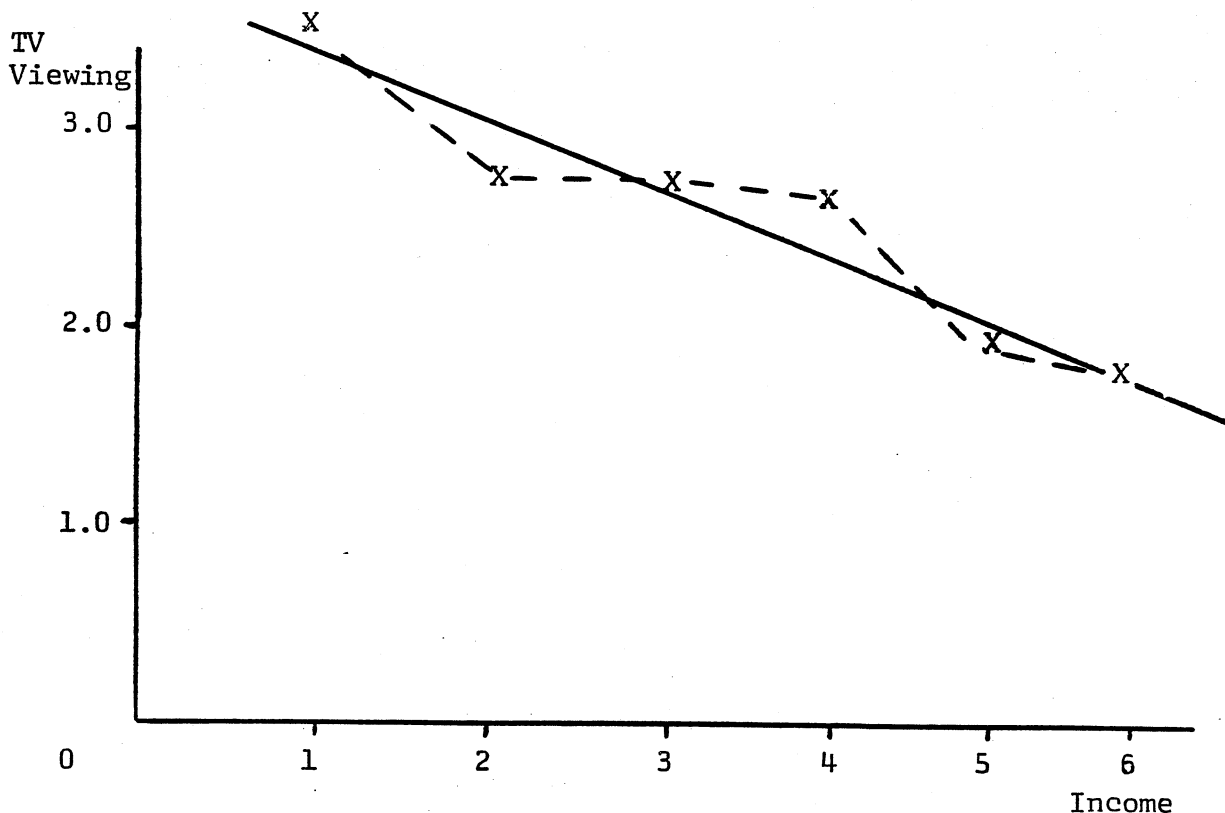


Figure 2. TV viewing vs. income.

TABLE 1  
TV Viewing vs. Income ANOVA

Source	Sum of Squares	D.F.
Between	145.909	
Explained by Regression	136.038	1
Unexplained by Regression	18.871	4
Within	2679.218	507
Total	2834.127	512
	$R^2 = .048$	
	$PR^2 = .878$	

F test gives:  $F = \frac{18.871/4}{2679.218/507} = 0.9$ , which is not significant, and hence the relationship, as thought, is essentially linear.

#### Value of $PR^2$

The value of the  $PR^2$  measure, therefore, is that it indicates whether a linear model appropriately expresses the relationship among a dependent and a set of independent variables. It does not, however, indicate the relative predictive value of independent variables. For example, a  $PR^2$  of .98 could result from variables which are not as useful in prediction as another set with a lower  $PR^2$ . Hence  $PR^2$  merely indicates how well a linear and additive (non-interactive) model explains the effect of a set of discrete independent variables on a dependent variable.

#### Conclusion

A variety of reasons exist which make the correspondence between goodness of fit and model usefulness less than perfect. This does not mean that goodness of fit measures are of little value in evaluating or developing a model. It does mean, however, that care is required in using goodness of fit measures for model evaluation. This suggests:

1. Be more explicit about exactly what the model is. This would include reporting post-hoc models as such.
2. Consider exactly how high a goodness of fit measure is reasonable given random behavior and measurement error. This means being suspicious of tautologies when  $R^2$ 's get above .6 in cross-sectional regressions of survey data.
3. Attempt whenever possible to utilize predictive testing to examine model validity.
4. Consider using special goodness of fit measures such as  $PR^2$  for examining the usefulness of stochastic models.



In summary, goodness of fit measures are useful tools in the hands of sophisticated researchers. A high goodness of fit, however, is neither a sufficient nor even a necessary condition for model usefulness.

## FOOTNOTE

1. Donald R. Lehmann is Associate Professor at Columbia University Graduate School of Business.

## REFERENCES

- Basman, R. L. On the application of the identifiability test statistic in predictive testing of explanatory economic models. The Econometric Annual of the Indian Economic Journal, 1965, 13, 387-423.
- Bass, F. M. Unexplained variance in studies of consumer behavior. In J. U. Farley and J. A. Howard (Eds.), Controlling noise in marketing data. Forthcoming.
- Bass, F. M. The theory of stochastic preference and brand switching. Journal of Marketing Research, 1974, 11, 1-20.
- Bass, F. M. & Clarke, D. G. Testing distributed lag models of advertising effect. Journal of Marketing Research, 1972, 9, 298-308.
- Bass, F. M., Tigert, D. J., & Lonsdale, R. T. Market segmentation--Group versus individual behavior. Journal of Marketing Research, 1968, 5, 264-270.
- Beckwith, N. E. & Lehmann, D. R. The importance of differential weights in multiple attribute models of consumer attitude. Journal of Marketing Research, 1973, 10, 141-145.
- Blalock, H. M. Causal inferences in nonexperimental research. New York: W. W. Norton, 1964.
- Cohen, J. B., Fishbein, M., & Ahtola, O. T. The nature and uses of expectancy-value models in consumer attitude research. Journal of Marketing Research, 1972, 9, 456-460.
- Hulbert, J. & Lehmann, D. R. Assessing the importance of the sources of error in structured survey data. In J. U. Farley & J. A. Howard (Eds.), Controlling noise in marketing data. Forthcoming.
- Lehmann, D. R. An index for assessing the appropriateness of a linear regression model. Columbia University Graduate School of Business, working paper, 1974.
- Lehmann, D. R. Preference among similar alternatives. Decision Sciences, 1972, 3, 64-82.
- Morrison, D. G. Evaluating market segmentation studies: The proper-

- ties of  $R^2$ . Management Science, 1973, 11, 1213-1221.
- Morrison, D. G. Regression with discrete random variables: the effect on  $R^2$ . Journal of Marketing Research, 1972a, 9, 338-340.
- Morrison, D. G. Upper bounds for correlations between binary outcomes and probabilistic predictions. Journal of the American Statistical Association, 1972b, 67, 68-70.
- Morrison, D. G. On the interpretation of discriminant analysis. Journal of Marketing Research, 1969, 6, 156-163.
- Stevens, S. S. On the theory of scales of measurement. Science, 1946, 103, 677-680.
- Taylor, J. & Gutman, J. A reinterpretation of Farley and Ring's test of the Howard-Sheth model of buyer behavior. In S. Ward and P. Wright (Eds.), Advances in consumer research. Vol. 1. Urbana, Ill.: Association for Consumer Research, 1973. Pp. 438-446.
- Wilkie, W. L. & Pessemier, E. A. Issues in marketing's use of multi-attribute attitude models. Journal of Marketing Research, 1973, 10, 428-441.



## VALIDITY

Larry Light<sup>1</sup> and Fran Kahn<sup>2</sup>  
 Batten, Barton, Durstine & Osborn, Inc.

"That's interesting, but is it valid"? A statement often expressed by marketers. What does the marketer mean by "Valid"? he means, "is it right"? Will this technique produce information that leads to correct decisions? The key to the marketer's concept of validity depends on the decision the marketer plans to make. How will the technique be used? Let us note that the marketer may be concerned with predictive validity. That is, how good is this research technique in predicting the probable outcome of some decision? He may be concerned with concurrent validity. Does this measurement correlate with what is going on at this time? In this paper, we will examine predictive validity. There are several important points that are worth noting if we researchers are to succeed in improving the validity (predictive) of our research procedures.

We cannot avoid the fact that we are trying to predict the effect of alternative decisions. This is an important assumption. First, if there are no alternatives, who needs research to help make a decision? Second, the effects to be predicted must be specified in measurable terms. Third, the effects of the decisions may be masked by a complex mixture of contaminating variables. Careful research design is a must.

Trying to predict outcomes, is the same as answering the question, "What will happen if \_\_\_\_\_"?

This is a reasonable test of validity, but we better be careful. There is a lot of momentum in the marketplace; most decisions about individual components of the marketing mix may have a relatively small effect: the effects of marketing decisions may take a lot of time to show up. We can easily be deceived. All we have to do is predict no effect. If we don't take care to take precise measurements (which may be expensive), we can be misled into believing our techniques are valid. To test the predictive validity of a research technique, we must predict changes in the criterion measure. The measure of predictive validity is the correlation of predicted changes with observed changes in the field criterion.

This kind of research is difficult, takes time, and costs money. This is unfortunate. But, the facts are that we need to do validation of this kind or we may be misleading a lot of people.

Since answering "What would happen if \_\_\_\_\_" is difficult, many researchers have adopted another procedure which they call validation. They ask, "does this technique correlate with some other technique"?

One of the two techniques is commonly accepted, or has face validity. This test may, of course, be misleading, since two techniques may correlate, yet both may be invalid predictors.

Assume, however, that we have properly validated a research test. Now, we want to validate another.

Let's look at a common experiment. Suppose we were trying to determine the validity of a new commercial testing system. A frequently adopted procedure is to take a set of "high-low" pairs (obtained from testing on the accepted criterion system). These pairs of commercials are tested on the new system. The results look like this.

	PAIR 1		PAIR 2		PAIR 3		PAIR 4		PAIR 5	
	A	B	A	B	A	B	A	B	A	B
I CRITERION SYSTEM	37	28	29	21	22	13	14	8	27	12
II NEW SYSTEM	45	35	66	55	59	51	50	38	49	36

Let us compare the results from system I and system II.

The new technique does seem valid. In five out of five cases the better commercial in a pair (as determined by system I) was correctly identified by system II. But is system II really valid?

It turns out that past experience with system I and system II permits us to classify the observed scores as "good", "average", "poor". So, we can re-examine the data in table I.

	PAIR 1		PAIR 2		PAIR 3		PAIR 4		PAIR 5	
	A	B	A	B	A	B	A	B	A	B
I CRITERION SYSTEM	G	G	G	A	A	P	P	P	G	P
II NEW SYSTEM	A	P	G	G	G	A	A	P	A	P

Now we see a completely different picture. If a reasonable purpose of a commercial testing system is to identify potentially effective advertising, how good is the new system? We have four "good" commercials (according to the criterion system). The new system correctly identified only one. Furthermore, the new system mis-identified two "average" commercials as good ones. This does not seem to be very valid, does it?

Yet, this is the same technique that previously seemed so valid. Is the test valid or not? The answer is unfortunately, both yes and no.

The new test is valid, if the purpose of the test is to compare alternatives. Assuming that a marketer has a pair of commercials and must run one of the two. The new system is valid for this purpose. The new system is acceptable for comparing alternatives.

But, what if the marketer wants to evaluate a given commercial? He wants to know if it is good. The new system does not seem to be acceptable for this purpose.

So, the researcher is faced with a dilemma. Should he recommend the use of the new system? The answer is that the new testing system may be recommended as a technique for screening down a set of alternatives, but there is no basis for recommending the new system for purposes of evaluation.

Valid? For what purpose? Screening? Or Evaluation? The proper answer to the validity question depends on a proper answer to these questions.

This example of a validation study, assumed that we had some knowledge of the predictive validity of a criterion system. Unfortunately, this is not generally the case. The reason usually given for the lack of predictive validity data is that it costs money.

The burden of this cost has traditionally been placed on the shoulders of the "buyers" of testing techniques. They have had little choice. Purveyors of new techniques have shown little or no willingness to invest in properly designed validation studies. So the buyer, the user, must either do his own validation study or have lots of faith. The "burden of proof" that a testing system is valid seems reasonably to rest with the seller, not with the buyer.

Sure it costs money to validate. But, this cost can be amortized over future uses of the testing system. More importantly, we should not ask, "what does validation cost?" We should ask, "what is validation worth?"

It's worth a lot to companies like Du Pont, Scott Paper, General Foods, Pillsbury, and others.

It's worth a great deal to know whether a technique is valid. It may be worth even more to know that a technique you may wish to use is invalid. Validity----- what's it worth to you?



## DISCUSSION: VALIDITY PROCEDURES IN CONSUMER BEHAVIOR RESEARCH

Albert C. Rohloff<sup>1</sup>  
Market Science Associates

The validity we look for is the validity we find. The four papers presented here examined many kinds of validity related to data, measurements, models, goodness of fit, constructs. We can infer from these papers, and our own experience, that a researcher will emphasize certain kinds of validation over others.

A specific example illustrates this: Suppose data is available that shows how soon consumers make their first purchase of a new product. This data is presented as a cumulative trial curve. The experienced eye of the analyst indicates that an exponential growth curve would fit this data. Goodness of fit criteria demonstrate that this model fits quite well, not only historical data of the company but data for introductions of other new brands of the company as they occur. The model has predictive validity. The model has been validated in the market place.

A marketing manager may take a different view: It is too good to be true (valid). The market I know is more complex than this. What is wrong here? How can the marketing plan for the next new brand introduction be changed to make this model fit poorly? Notice what the manager is suggesting is an invalidation procedure.

Experience has shown that when a simple exponential growth curve fits trial curves well, it is usually symptomatic of an inefficient (even ineffective) marketing plan. The better the model fits the poorer the plan tends to be!

This example illustrates my opening proposition: The validity we look for is the validity we find.

Lehmann appears to be placing "the predictive model" in a central position. This approach can be limiting if the predictive model becomes an end in itself. Pessemier states: "The commercial researcher's most important problem is not simply to employ valid data and models...it is more difficult...He must find valid relationships that are managerially important."

How can this "managerial" validity be increased? I suggest that if managerial validity is what we are looking for, it is the managerial validity we will find.

An excellent starting point for achieving managerial validity for consumer behavior research is to understand what are the basic beliefs and theories the manager has about consumer behavior. Some of his beliefs can be inferred from the way he markets a brand. When research is keyed to managerial validity, it is natural to identify which beliefs about consumer behavior are most critical and then proceed to test them against alternative viewpoints that may be more valid.

But consumer behavior is, in part, determined by the manager's beliefs. Note that as long as the company introduced new brands in the same way, the feedback on trial purchases was consistent. When an innovative plan was used, consumer behavior changed.

In the appliance example Pessemier describes the creative interpretation of



relationships, noting that other research and consumer behavior is also brought in. Often the research is seen as providing the facts and the manager the creative interpretation. As the researcher is more closely integrated in the marketing planning team, he has more opportunity to contribute towards "managerial validity".

Pessemier sets up a dichotomy of academic research and commercial research, seeming to say that the validation criteria for academic research are less demanding. There is a greater demand on academic research for valid consumer behavior theory. The example given demonstrates a serious limitation when the predictive model is held up as the criterion of ultimate validity for consumer behavior research. A model is not a comprehensive theory.

For both the commercial and academic researcher the validity he looks for is the validity he finds. Both can profit by taking a second look.

#### FOOTNOTE

1. Albert C. Rohloff is Senior Vice-President, Market Science Associates, New York.

## VALIDITY PROCEDURES IN CONSUMER RESEARCH: A PERSPECTIVE

Albert V. Bruno  
University of Santa Clara

This session was concerned with validity procedures in consumer research. Kerlinger has described the subject as "complex, controversial, and peculiarly important" (Kerlinger, 1964). Our session has identified additional descriptions such as elusive, expedient, and expensive.

Papers concerned directly or indirectly with validity issues are commonplace. Virtually every researcher purports to be interested in various aspects of validity.

Much of the research reported in journals relevant to our interests contains treatment of validity considerations which is vague and poorly defined, ill conceived and often constrained by the data on which the constructs are built, and not clearly identified as being cost justifiable -- validity procedures are virtually incidental to the research process.

The papers in this session represent the viewpoints and interests of a diverse set of individuals. Pessemier's paper "Incorporating Tests of Data and Model Validity in Commercial and Academic Research" discusses the need for early consideration of validity issues, incorporating validity tests in the research effort and reporting the validity of the measures and models along with other research results. Juster's paper describes "Validity Procedures at the Survey Center" as a guide for researchers. Lehmann's paper, "Validity and Goodness of Fit in Data Analysis," describes the relationship between the usefulness of a hypothesized relationship between two or more variables and the goodness of fit which results from an empirical investigation of the hypothesis. Light in "Validity -- What's It Worth" investigates the value of validity from an ad agency's point of view.

The aforementioned textual material and comments by the discussant, Rohloff, focuses on several important questions for researchers:

- 1) Validity and validity procedures. The content of the papers, the comments of the discussants, and the observations of several members of the audience demonstrated that we are not in agreement as to the nature of validity or its importance in the scientific research process. If consumer research is to rise to its appropriate position in the hierarchy of scientific inquiry, we must develop some consensus as to what validity is and why it is important.
- 2) Managerial vs. scientific validity. A number of researchers appear prepared to distinguish between managerial significance and significance based on statistically significant differences. If the function of our research is to aid management in making better decisions, can these distinctions exist? If there is some legitimacy in distinguishing between the two, under what circumstances and conditions can the case be made?
- 3) Cost/benefits of validity. Much of the treatment of validity and validity procedures has dealt with validity independent of its cost. Are all validity activities cost justifiable? Under what circumstances can some validity activities be eliminated?

## Recommendations

The resolution of these and related issues can be achieved only if a consistent, objective methodology becomes commonplace. Toward that end, it would be useful to develop consistency in reporting of research results. The following recommendations are made:

1) Identification of "criterion" variables. Researchers should clearly identify the nature of the "ultimate" criterion variable(s) used for evaluating the validity of the measure under study. Conclusions and implications of research studies can not be derived unless some prior evaluation, and hopefully, as Heeler and Ray have recommended (Heeler and Ray, 1972), validation of the criterion variable has been performed. An example may illustrate this point. The Likert-scale approach to identifying the brand preference of consumers may have a conceptual foundation and substantial empirical research support; consequently, it may be well accepted as an "ultimate" criterion variable. Subsequent research which utilizes this form of brand preference (e.g., attitude/behavioral studies) will thus be on some common footing among groups of researchers. However, if a researcher reports results which utilize a new, unsubstantiated form of brand preference (e.g., constant sum score preference) in an attitude/behavioral research project, it may be difficult to draw useful conclusions from the researcher's efforts. Some determination of the appropriateness or legitimacy of the criterion variable should be included as an integral portion of research studies, either by linkage to previous research or, at least, by intellectual exposition.

2) Distribution of responses. The mean, standard deviation, and especially, the distribution of responses of both the criterion variables and the measures of interest should be reported or at least discussed by the researchers. Researchers should take precautions to insure that no "abnormal" distributions are "creating" results. For example, if there is no spread over the response categories (i.e., low variance for the question), then high correlations can be the result of consistent responses from a small number of respondents. As Ferber has suggested, "the existence of a relationship between the attributes is in doubt until the presence of a possible fluke is investigated and further studies are made" (Ferber, 1949). Although rather pedestrian, a scatter diagram would provide extremely useful information.

3) Complementary analyses. In addition to the use of product moment correlations, it is strongly recommended that other methods of validation be employed. Cross classification analysis can provide additional focus on the nature of the relationships identified by the correlations. Moreover, since predictive validity is often the ultimate goal of marketing research, the use of "hit or miss" tables and "confusion matrices" would provide strong tests of the predictive validity of the measures under consideration.

4) Application of validation procedures. There are different types of validity which have been described adequately elsewhere by a number of authors (for example, see Bruno and Pessemier, 1972; Ferber, 1949). A hierarchical ordering of validity procedures in marketing research should descend from predictive validity (at least as long as "goodness of fit" serves as the conventional measure of effectiveness). Marketing researchers who are concerned with prediction (as opposed to being concerned with theory development, for example) should be concerned with identifying measures with high predictive validity rather than with identifying measures that are reliable, have content and construct validity, but have low predictive validity. If a researcher were

interested in identifying brand preference measures that predict subsequent choice of brands, then the ultimate "test" of the usefulness of the brand preference measures is their ability to correctly predict brand choices. Other types of validity are of secondary importance.

Too many marketing researchers have been guilty of utilizing less desirable forms of validity because of cost or time limitations or because their data were not in the appropriate form. Occasionally, researchers even appear to have a misconception of what is being researched. The author of a recent article entitled "Linear Attitude Models: A Study of Predictive Ability" purports to examine the question of predictive validity by correlating actual and predicted preference (Churchill, 1972).

5) Sensitivity to changes in assumptions. The validation procedures may be dependent to a degree on explicit or implicit assumptions by the researcher. Efforts should be made to isolate the degree of dependency and determine if it is acceptable. Perhaps a study of the dependency through some form of sensitivity analysis would be appropriate. Recent investigations on the question of the optimal number of response categories of Likert-type scales is an example of this approach (Green and Rao, 1970; Lehmann and Hulbert, 1972). Certainly, the degree of predictive validity (if correlation coefficients are utilized as criteria) of some measures is partially dependent upon the number of response categories.

6) Need for synthesis and consolidation. There is a substantial need for synthesis, consolidation, and classification as a prerequisite to the development of scientifically satisfactory and managerially useful attitudinal/behavioral methodology. The time has come for editors and reviewers to insist on more systematic, scientific, and objective investigation of measure validation.

#### FOOTNOTE

- 1) Albert V. Bruno is an Associate Professor at the Graduate School of Business, University of Santa Clara.

#### REFERENCES

- Bruno, A. V., & Pessemier, E. A. An empirical investigation of the concurrent validity of selected attitude and activity measures. In M. Venkatesan (Ed.), *Proceedings of the 3rd Annual Conference of the Association for Consumer Research*, University of Chicago, November, 1972.
- Churchill, G. A. *Linear attitude models: A study of predictive ability*, Journal of Marketing Research, 1972, 9, 423-426.
- Ferber, R. Statistical techniques in market research. New York: McGraw-Hill Book Company, Inc., 1949.
- Green, P. E., & Rao, V. D. Rating scales and information recovery--how many scales and response categories to use? Journal of Marketing, 1970, 34(1), 33-39.
- Heeler, R. M., & Ray, M. L. Measure validation in marketing. Journal of Marketing Research, 1972, 9, 361-370.

Kerlinger, F. N. Foundations of behavioral research, New York: Holt, Rinehart, and Winston, 1964.

Lehmann, D. R., & Hulbert, J. Are three-point scales always good enough? Journal of Marketing Research, 1972, 9, 444-446.

# OVERVIEW OF POLITICAL CANDIDATE MARKETING<sup>1</sup>

Philip Kotler<sup>2</sup>  
Northwestern University

Political campaigns have increasingly been compared to marketing campaigns in which the candidate puts himself on the voters' market and uses modern marketing techniques, particularly marketing research and commercial advertising, to maximize voter "purchase." It is argued here that the very essence of a candidate's interface with the voters is a marketing one and always has been. Alleged differences between commercial marketing and political marketing are shown to be overstated.

The American public is treated every few years to the illusion that they run their country by making a choice among the men who want to run their country for them. During the 1972 presidential election year the candidates for various national, state, and local offices managed to somehow spend over \$400 million in less than two months on their campaigns (Weiss, 1973). The presidential candidates themselves spent 25 percent of this sum, or \$98 million (Advertising Age, 1973). Of this, the Republican candidate spent 61 percent of the total, or \$60.2 million. Needless to say, he won the election.

Political campaigns have increasingly been compared to marketing campaigns in which the candidate puts himself in the vote market and uses modern marketing techniques, particularly marketing research and commercial advertising, to maximize voter "purchase." The marketing analogy is more than coincidental. It is argued here that the very essence of a candidate's interface with the voters is a marketing one, not only in recent times but far back into the past. Candidates seeking to win elections cannot avoid marketing themselves. The only question is how to do it effectively and responsibly.

Interest in the marketing aspects of elections has been stimulated to a large extent by the spectacular growth in political advertising. There has also been a substantial growth of scientific opinion polling (i.e., marketing research), computer analysis of voting patterns (i.e., sales analysis), and professional campaign management firms (i.e., marketing organizations). The subtleties of the marketing approach go beyond the rising expenditure levels and the use of certain information and planning approaches. They are delineated in a series of popular books (White, 1961; McGinness, 1969). In a quieter way, several scholarly works have also noted the marketing character of political elections (Kelly, 1956; Glick, 1967; Nimmo, 1970).

It would be a gross mistake to think that election campaigns have taken on a marketing character only in recent years. Campaigning has always had a marketing character. Prior to the new methodology, candidates sought office through the handshake, baby kissing, teas, and speechmaking. They still use these methods. The "new methodology" is not the introduction of marketing methods into politics but rather an increased sophistication and acceleration of their use. According to Glick (1960, p. 1):

The personal handshake, the local fund-raising dinner, the neighborhood tea, the rally, the precinct captain

and the car pool to the polls are still very much with us...the new campaign has provided a carefully coordinated strategic framework within which the traditional activities are carried out in keeping with a Master Plan. It centers on a shift from the candidate-controlled, loosely-knit, often haphazard "play-it-by-ear" approach to that of a precise, centralized "team" strategy for winning or keeping office. Its hallmarks include the formal strategic blueprint, the coordinated use of specialized propaganda skills, and a more subtle approach to opinion measurement and manipulation. And, though there is a world of difference between selling a candidate and merchandising soap or razor blades, some of the attributes of commercial advertising have been grafted onto the political process.

Nimmo (1970, pp. 67-68) takes a cynical view of this development:

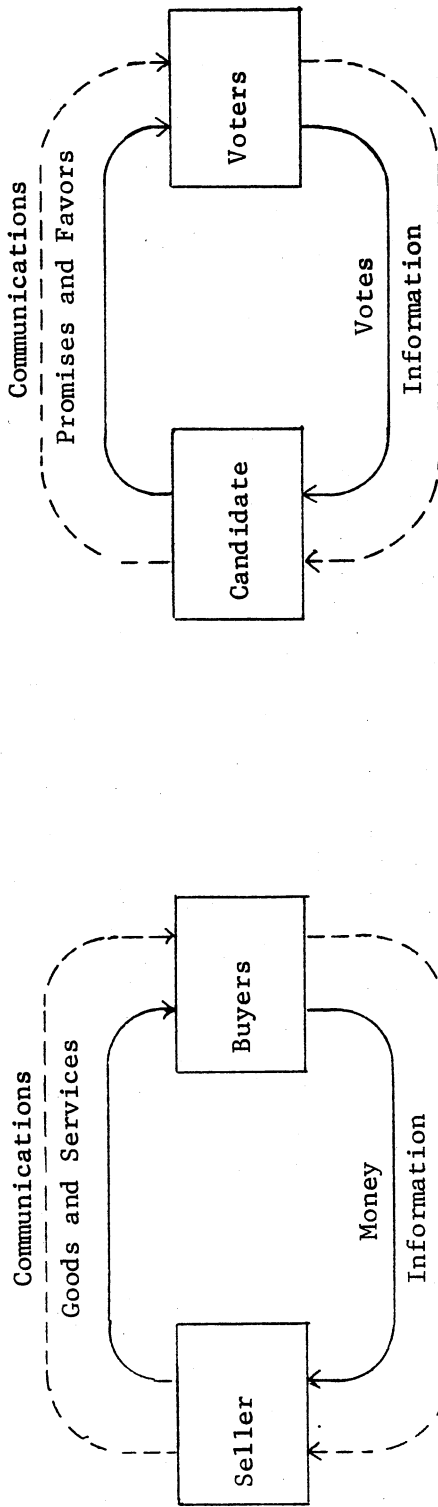
In screening potential candidates the mercenaries have given a new definition to the notion of "availability"; the marketable candidate is selected on the basis of his brand name, his capacity to trigger an emotional response from the electorate, his skill in using mass media, and his ability to "project." Analysis of social problems and issues yields to parroting of themes; televised debates between contenders produce meaningless confrontations rather than rational discussion. Negotiations with party politicians assume the form of "out-of-town tryouts"; primary elections are approached as "presale" campaigns; and general elections emerge as the "Giant Sweepstakes." In the end one candidate owes his election not to party but to his personal organization of paid and voluntary workers; once elected he responds not to party programs, but to the interests also represented by the professionals.

The major fault with Nimmo's observation is that it takes on a moral, judgmental tone. It implies that something is happening to political contests that is called marketing and it is bad. It fails to recognize that the marketing problem exists no matter what means or style of marketing is used. In fact, marketing styles vary from product to product and time to time; but the marketing problem is always present.

### The Anatomy of Political Marketing

Figure 1 presents two maps comparing business and political marketing. The business marketing map shows a seller dispatching goods, services, and communications to the market; in return he receives dollars and information. The inner loop is an exchange of money for goods; the outer loop is a flow of information. The political marketing map shows a political candidate dispatching specific promises and favors to a set of voters in exchange for their votes. He uses general communications to convey these and gathers voter information to plan next period's marketing effort. Schematically, the structural processes of business marketing and political marketing are basically the same. Both can be analyzed in terms of exchange theory.

Figure 2 shows the political candidate's marketing problem in greater detail. (1) The environment which defines the salient issues and oppor-



A. Business Marketing

B. Political Marketing

Figure 1. Business and Political Marketing Compared



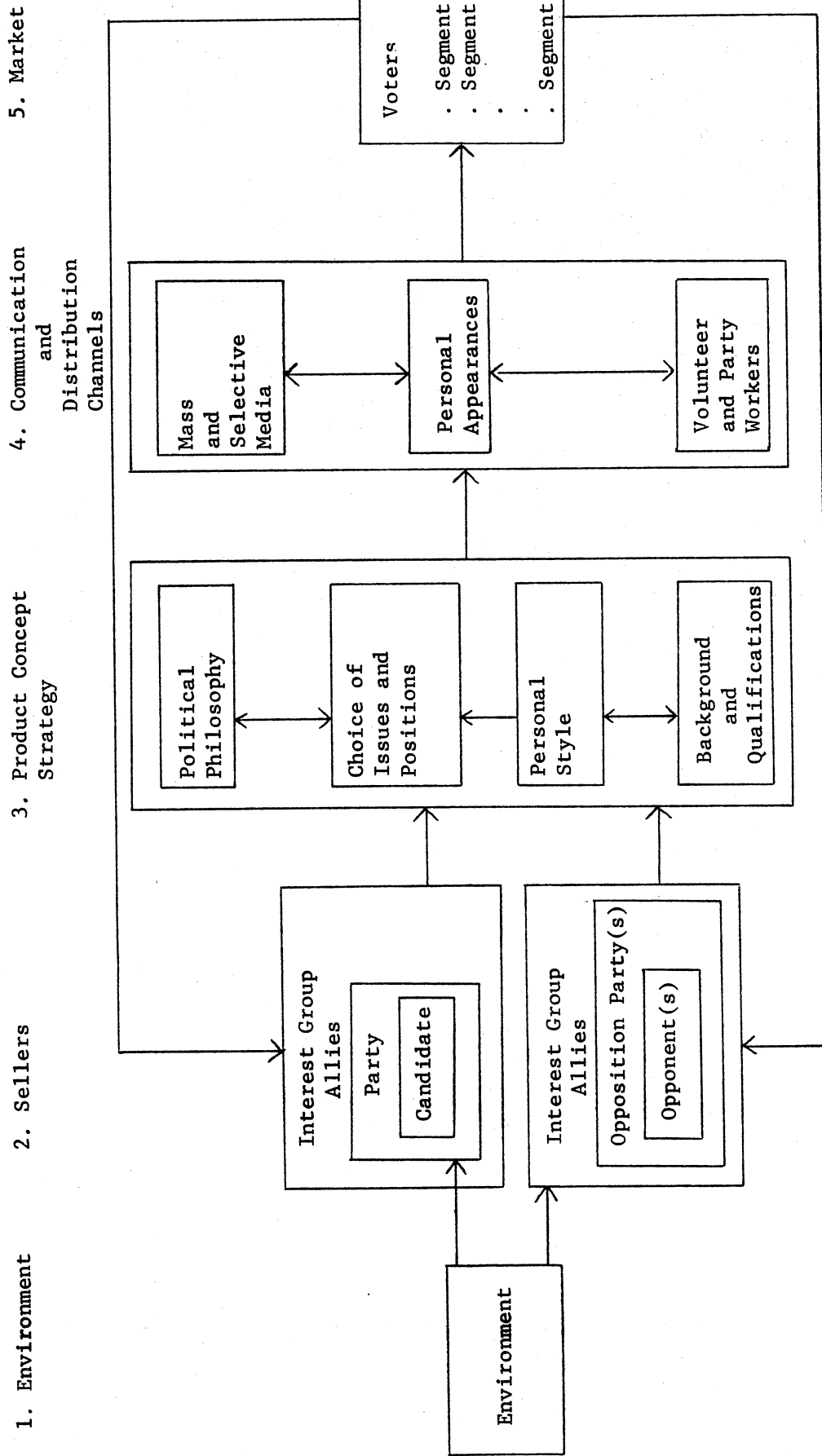


Figure 2. A Comprehensive Political Marketing Map

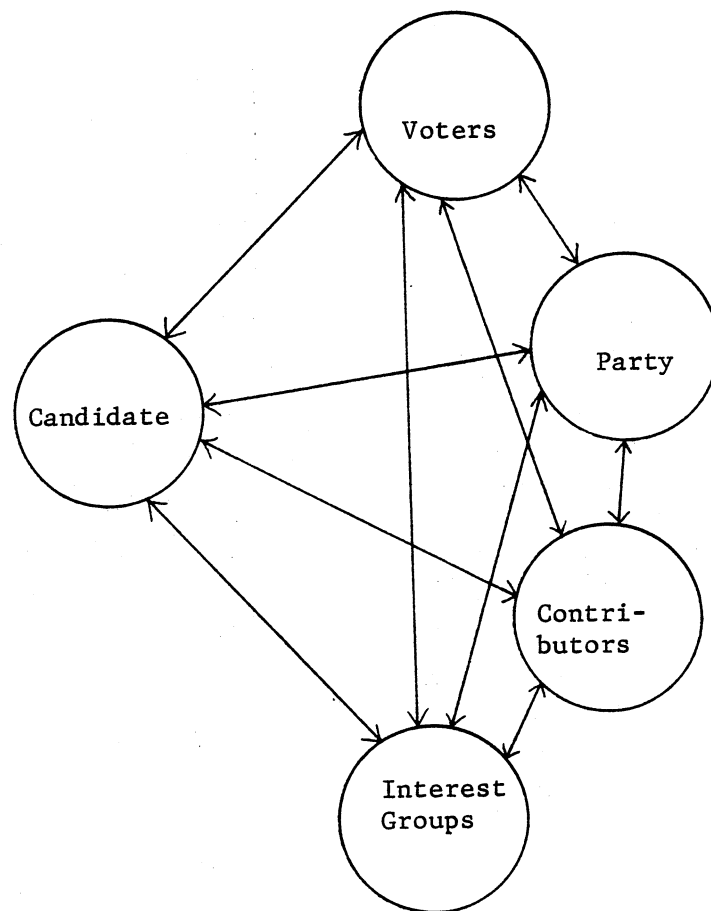


Figure 3. Four Markets Faced by the Candidate

tunities for the candidates is shown at the far left. (2) The candidates, their parties and their interest group alliances are the sellers. (3) Each candidate develops a product concept that he believes is merchandisable to the voters. The concept is built on a mixture of political philosophy, stands on particular issues, personal style, and background qualifications. (4) Each candidate seeks to reach the voters through three major distribution channels: mass and selective media, personal appearances, and volunteer and party workers. These channels interact, e.g., a personal appearance reaches an immediate audience and also a larger audience through mass media coverage. (5) All of these efforts are adjusted for different voter segments, and the results are continuously reviewed for further campaign modification.

The candidate must not only develop a marketing strategy calculated to win the support of voters but also of the party, contributors, and interest groups. Figure 3 shows the four markets he faces. The interactions among these markets are complex and the candidate cannot afford to formulate his marketing strategy simply on the basis of voter market response. For example, taking a strong stand as an anti-machine candidate will gain voters' votes but hurt party support and some contributions. Furthermore, each stand has distributive effects within a particular category. Taking a pro-labor stand increases the contributions from labor and reduces the contributions from business. Thus, political marketing strategy cannot be developed by simply calculating the distributive effects within the voters' market. Similarly, business marketing strategy aiming at building buyer support cannot ignore the impact on dealers, stockholders, government, and competitors.

#### Alleged Differences Between Political and Business Marketing

Certain differences between political marketing and business marketing are alleged to exist in the public mind. They relate to characteristics of the product, buyers and sellers.

1. Any specific commercial product, such as a can of beans or a ton of steel, is relatively fixed in its characteristics at a given point in time. The political candidate, on the other hand, is more variable. For one thing, the political candidate can talk back.

Comment. The variable nature of the political product is matched in the commercial world by services. Services are inseparable from the people who render them. The housewife can testify that her hairdresser can talk back, that his hairdos vary in quality, and so on. Physical products, too, can be changed through product reformulation, sizing, or packaging.

2. It is held that the political candidate cannot be as thoroughly formulated for the market's needs as can physical products or services. For example, new foods and soap products can be formulated to meet specific market wants. But a given political candidate cannot be varied freely in the same way. He has a history and fairly set personality. It is not easy or even possible to remake a humorless old candidate into a vigorous young one.

Comment. The freedom of manufacturers to alter the character of some commercial products is also quite limited. Steel is steel and salt is salt. On the other hand, a political image can be changed to some extent. Richard Nixon in successive campaigns has taken on the image of an anti-Communist, a statesman, and most recently, a man-of-action. There are limits, of course, and organizations often choose to launch brand new products rather than to do the job of repairing the old one.

3. Business products are normally available for purchase any time at the discretion of the buyer. Political products, however, are only "put on sale" every few years.

Comment. There are instances of economic products that buyers can only buy at certain times. One can buy a Rembrandt only when it is put up for sale or auction. One can enroll in a college only during certain times of the year. Many government contracts carry announced dates for bidding. Purchase frequency is not a basis for distinguishing between commercial and non-commercial products.

4. The buyer of a commercial product or service usually expects personal benefits enjoyed within a reasonable time period. Many citizen voters do not expect to accrue any personal or early benefits from their act of voting.

Comment. There are various commercial products and services that do not appear to give personal or short-run benefits, that people nevertheless buy. Examples are insurance, wills and estate plans, and so on. People also contribute to various charitable causes from which they do not anticipate personal benefits. On the other hand, many voters get quite involved in some election contests and act as if they anticipate personal benefits. The charismatic candidate is someone who gives a great number of voters the feeling that they will personally benefit through the candidate's election.

5. Buyers of commercial products and services are used to hard marketing tactics whereas voters do not expect, and somewhat resent, hard marketing tactics in the political area. A political candidate who would offer trading stamps or who overdoes hard-sell advertising would be taking great risks.

Comment. Hard-sell marketing tactics are characteristic of certain goods such as automobiles, cigarettes, soap and cosmetics. The marketing of earth moving equipment, computers and airplanes is conducted much more on the rational merits of the product and the company's reputation for service and reliability. There is nothing about marketing method that requires hard-sell tactics.

6. The messages reaching the public about a commercial product are largely marketer-controlled, through paid advertising. The media rarely feature or comment on a brand of beans or toothpaste. On the other hand, the messages reaching the public about a political candidate are largely developed by the news media. As a result, the political candidate finds it necessary to market himself as much to the press as to the ultimate public.

Comment. It is true that the press takes an active role in commenting and interpreting political candidates to the public. This makes the political candidate's marketing task easier in some ways (he gets more free publicity) and harder in other ways (he has less control over what they say).

7. A business firm succeeds if it obtains any market share that yields a good return on its investment; the political candidate succeeds only if he obtains a plurality of the votes, that is, the largest market share.

Comment. There are business markets, too, where the seller either gets all or none of the business. For example, an airframe manufacturer who bids for a government contract either wins or loses. The criteria of what constitutes a viable market share makes little difference to whether marketing planning and strategy are useful.

8. The aims and means of the business seller and the political candidate are different. The business seller is seeking profits. The political candidate is seeking power. The business firm tries to secure more profits through creating satisfied customers. The political candidate does not as clearly try to secure more power by creating satisfied citizens.

Comment. Business firms actually pursue multiple objectives, as do political candidates. There are, in fact, business firms that pursue power and political candidates who seek profits. Furthermore, business firms and political candidates can choose from a range of philosophies on which to base their marketing. There are politicians who aim at producing satisfied citizens; and business firms that aim at quick profits.

#### The Marketing Problems of the Political Aspirant

Assume that a candidate has decided to enter politics and his ultimate goal is to achieve an elective office. At the beginning, he is an unknown product. The office seeker must put himself on a market, the voters' market. He has to go through many of the steps that occur in product marketing: develop a personality (brand image), get the approval of an organization (company image), enter a primary election (market test), carry out a vigorous campaign (advertising and distribution), get elected (market share) and stay in office (repeat sales).

Specifically, he must solve four successive problems in the achievement of a successful political career.

First, he must join a political organization and become known. He will want to develop a political style that will earn respect and leadership in his party. He knows that this means finding out what the members of the political organization appear to want from the political process and the extent that he can appear to be instrumental in their desires.

Secondly, he must eventually exhibit an interest in becoming his party's candidate in an upcoming election. He must fraternize with the leaders and attempt to get their backing. He must enter a primary election and win the support of the party's voters.

Thirdly, if he wins the primary, he will have to go before the voters in the general election. He will have to make important decisions on campaign strategy, including issues, advertising, appearances, and funding. He will face a problem in voter analysis, choosing targets, allocating resources, and timing them for maximum impact.

Fourthly, if he is elected, he must do the kind of job in office that will get him reelected. This will be a function of the organization he builds, the positions he takes, and the rhetoric he uses.

At each stage, the political aspirant must have a good understanding of his market's needs, perceptions, and preferences. He must be guided by reliable and valid models of consumer behavior.

## FOOTNOTES

1. An expanded version of this paper will appear in Philip Kotler, Marketing For Nonprofit Organizations (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1975), Chapter 19.
2. Philip Kotler is Harold T. Martin Professor of Marketing, Graduate School of Management, Northwestern University, Evanston, Illinois 60201.

## REFERENCES

- Glick, E. The new methodology. Washington, D.C.: American Institute for Political Communication, 1967.
- Kelly, Stanley, Professional public relations and political power. Baltimore: Johns Hopkins Press, 1956.
- McGinness, Joe. The selling of the president. New York: Trident Press, 1969.
- Nimmo, Dan. The political persuaders. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1970.
- "Slush and "shush." Advertising Age, October 8, 1973, 14.
- Weiss, E.G. Political advertising blackens the other eye of the ad business. Advertising Age, February 12, 1973, 35-38.
- White, Theodore. The making of the president 1960. New York: Atheneum House, 1961.



## IMAGES AND VOTERS' DECISION-MAKING PROCESSES

Dan Nimmo  
University of Tennessee

The language of contemporary political campaigns frequently likens the voter to a consumer in the political market place. This paper explores various facets of voters' political images as components of their decision-making processes as consumers. Using data from a nationwide survey conducted in the 1972 presidential election, it relates cognitive, affective, and conative aspects of voters' self-images, candidate images, and party images to voting behavior. Correlation, factor analytic, and stepwise regression techniques indicate that voters' affective responses to the major party candidates, rather than cognitive and conative considerations, were the key components in the electorate's 1972 decision-making processes.

During the last century the language of political campaigns possessed a martial tone. "Cadres" of party workers, obeying party "captains" organized in a chain of command, "drilled" phalanxes of "loyal" voters and scorned "traitors" and "turncoats." Shortly after the turn of the century, according to historian Richard Jensen (1968), a merchantilist style supplanted older political ways. In the new argot what was once a battlefield became a "market place," elections "sales campaigns," advertising the "pitch," and voters "consumers." Accepting more recent lexicon, this paper explores what considerations enter voters' decision-making processes when they act as consumers in the political market place.

To bridge the gap between what political scientists know about electoral behavior and the notion that voters are consumers, we use a concept in vogue among professional campaigners and journalists as well as political scientists and economists--i.e., that of image. Published popular and scholarly commentaries point to the importance of the images of political candidates to the electorate's voting decisions: In his widely read account of the 1968 presidential election McGinniss (1969) offered the gist of image politics, that "style becomes substance" where "the Medium is the message and the masseur gets the votes [p. 30]." A specialist in electronic politiking added to claims made for candidate images by a portrayal of the techniques of televised packaging of images (Wyckoff, 1968). A former Chairman of the Federal Communications Commission, Newton N. Minow (1973), has declared that "television's most significant political characteristic probably is its ability to present an image of a politician--providing an indication of his character and personality [p. 6]." And political scientists Natchez and Bupp (1968), drawing upon voting in presidential elections from 1952-1964, concluded that the "best single predictor of voting behavior . . . is candidate image [p. 411]."

An understanding of the relation between images and voters' decision-making processes requires, first, a definition of image and, second, a model of voting decisions. For the definition it is convenient to turn to Boulding's discussion in his provocative monograph (1956). Following Boulding we say that each person possesses an image, or series of images, of the world.



Each image consists of the person's subjective understanding of things--i.e., of what he or she believes to be true about something, likes or dislikes about it, and proposes to do about it (what social psychologists such as Smith, 1947, refer to as a person's cognitions, affects, and conations). This use of image parallels the definition of "brand image" in advertising and market research. Thus, Downing (1964) defines a brand image as "a constellation of feelings, ideas and beliefs associated with a brand by its users and non-users mainly as a result of experience of its advertising and performance [p. 14]." As with brand images, political images do not exist apart from the political objects (or their symbolic surrogates) that stimulate political thoughts, feelings, and inclinations. In sum, an image is a human construct imposed upon an array of perceived attributes projected by an object, event, or person. Thus, for instance, a candidate's "image" consists of how voters perceive him/her, perceptions based upon both the subjective appraisals made by the voters and the messages (utterances, attributes, qualities, etc.) transmitted by the candidate.

Assuming such a definition of image will serve us, what model of voting decisions will suffice? Numerous voting studies in recent decades suggest such a model (Kelley and Mirer, 1974) especially those of presidential elections carried out by the Survey Research Center/Center for Political Studies (SRC/CPS) of the University of Michigan. The first major volume in the Center's voting studies, prepared by Campbell, Gurin, and Miller (1954), was an extensive analysis of the 1952 presidential election. It identified major components of the voter's "motivation" to select one candidate over another. The authors pointed to three components of that motivation--the voter's party identification (whether he thought of himself as a Democrat, Republican, or Independent), orientation toward election issues, and orientation toward the competing candidates. Campbell, Gurin, and Miller demonstrated that these three forces directly related to candidate preference so that, for example, the stronger a person considered himself to be a Republican, and/or liked the Republican candidate, and/or preferred Republican issue postures, the more likely he would vote for Eisenhower. The less distinct his stand on any of these three factors, the more likely his choice of the opposition's party candidate (or, in many instances, the less likely he would vote at all).

Here, then, was an early statement of three key variables--party, candidate, and issue orientation--that enter the voting decision. In a second volume Campbell, Converse, Miller, and Stokes (1960) raised party identification to a prime position. Party identification--or, more aptly, the partisan self-image of the voter (Butler and Stokes, 1969)--acts as a predisposing attitude, usually formed early in life, and a perceptual screen influencing how each voter sees and evaluates candidates, issues, and political parties. In short, partisan self-images influence other political images. Moreover, partisan self-images are longterm influences on the voter farther removed from the voting act than other short-term influences such as the voter's images of the candidates and parties or issues of domestic and foreign policy (Campbell, Converse, Miller, & Stokes, 1966).

The question that concerns us within this context of longterm and short-term components is what is the relative impact of voters' images in making electoral decisions? Generally, as Natchez points out in his thoughtful review and critique (1970), published studies of the SRC indicate that "the most powerful components of the electoral decision process are those which capture attitudes toward the competing candidates [p. 575]." In the early presidential

elections following World War II domestic nor foreign policy issues had a great influence on election outcomes; also, attitudes toward the political parties and other political groups were not generally associated with shifts in voters' preferences between elections. Rather, candidate image was the primary motivator of voting decisions. Although the pattern has changed some, candidate images retain considerable potency. For example, using SRC/CPS data from the 1960 and 1964 presidential elections, Natchez and Bupp (1968) found that issues clearly had come to be of considerable influence on voting, but candidate image remained more important.

Along with Natchez and Bupp, studies of more recent presidential elections also indicate an increased role for issues in shaping voting behavior, yet these investigations confirm that candidate image remains a principal short-term force. In 1968, for instance, the Comparative State Election Project (Kovenock, Beardsley, & Prothro) focused upon voting behavior in Southern states (combining several statewide surveys with a nationwide sampling). By comparing respondents' positions on selected issues with their perceptions of the candidates' stands on the same issues, the analysis explained a considerable portion of the variation in voters' preferences for Nixon, Humphrey, and Wallace. Although issues were clearly important in 1968, however, a key component of the voting decision was the voters' images of the candidates' issue stands along with their perceptions of other candidate qualities. And, an analysis of the 1972 presidential election (Miller, Miller, Raine, & Brown, 1973) also found that how warmly or coldly voters evaluated the candidates and the perceptions of voters of the candidates' issue stands compared with their own positions were major factors in helping voters make up their minds.

Natchez and Bupp observed (1968) that the longterm component of voting--partisan identification--was, to be sure, "causally" associated with candidate image so that "people do not perceive candidates through neutral eyes" but "seem to take their focus from previously established identifications [p. 446]." Yet, partisans sometimes have more positive images of the opposition party's candidate than they do of their own. Boyd (1969) demonstrated, at least for the presidential elections of 1956-1964, that when partisans have images of the candidates that are inconsistent with their partisan self-images, this constitutes the major factor explaining their defection from party loyalty in voting. Using SRC/CPS data Boyd derived measures of consistency between the longterm component of a person's voting decision, party identification, and each of three short-term components--i.e., the voter's image of the candidates, of the political parties, and of selected policy issues. Boyd also calculated a measure of the degree to which voters of varying intensities of partisan self-image voted in each election in keeping with their partisanship or defected to the opposition. Comparing the defection rates of voters with the consistency of their partisan self-images and images of candidates, parties, and issues he found that images of the candidates was the principal statistical explanation for voting defection.

Not only does contemporary research indicate that candidate image is the most important short-term force contributing to partisan preference and defection, but it also suggests that how people perceive candidates is the principal determinant of whether they will split their tickets and how. In 1972 Miller and Jackson (1973) conducted a probability sampling of registered voters in three counties of southern Illinois. Their focus was upon the factors related to voters splitting their tickets between the two major political parties in contests for the presidency (Richard Nixon vs. George McGovern) and the governorship (Richard Ogilvie vs. Daniel Walker). Respondents evaluated each of the four candidates using various seven-point semantic

differential scales. With the same scales respondents also rated traits of their Ideal President and Ideal Governor. The researchers derived measures of the proximity between respondents' evaluations of each presidential candidate and their image of the Ideal President; they also calculated proximity measures for the gubernatorial candidates. Miller and Jackson found that candidate images (that is, how closely each voter's image of any candidate approximated the voter's image of the "ideal" office-holder) enabled them to predict straight and split-ticket voting with considerable success.

Running through the bulk of major recent research relating various components of voting decisions to electoral behavior, then, has been a view that "candidate images" (or, more accurately, voters' images of candidates) are powerful short-term forces in the electorate's decision-making. The remainder of this paper examines that possibility. Specifically it seeks to identify what dimensions in the relationship between voters' issue, candidate, and party images influence voting decisions--i.e., are beliefs about candidates/party/issues (cognitions), feelings (affects) toward these political objects, or inclinations (conations) regarding them most influential? Using data from the 1972 presidential election study of the CPS (based upon a nationwide probability sample) it is possible to measure five selected components of the voting decision--partisan self-images and ideological self-images of voters, issue orientations of the electorate, and voters' images of the candidates and of the political parties. By correlating and factor analyzing such measures and conducting a regression analysis we can describe the relative impact of each component upon voting.

Of the five components, three pertain to a voter's self-image; i.e., what kind of partisan the voter perceives he is, what kind of liberal or conservative he thinks he is, and what stands he takes on specific policy questions. Our interest in these is in the voter's perceptions of self, not of the candidates or the political parties. The SRC/CPS has been measuring the distribution of partisan self-images in America for more than two decades by asking cross-sections of American citizens a series of questions regarding whether they "usually think" of themselves as Republicans, Democrats, Independents, or something else. This practice in 1972 produced a seven-point scale of partisanship (ranging from a score of "1" for strong Democrat through "3" for a strict Independent to "7" for strong Republican). Also in the 1972 election survey respondents placed themselves on a seven-point scale measuring their ideological self-images: viz., extremely liberal, liberal, slightly liberal, moderate/middle of the road, slightly conservative, conservative, and extremely conservative. Finally, we use a measure for the voter's issue stands. Respondents in the 1972 survey, again using a seven-point scale, revealed their positions on several policy issues. We employ nine issues in this analysis covering a diverse gamut of social, economic, foreign policy, civil rights, and civil liberties matters: (1) whether the federal government should guarantee that every person has a job and a good standard of living, (2) whether income taxes should be progressive or everyone pay the same portion of their income regardless of how much they make, (3) the question of legalization of the use of marijuana, (4) busing to achieve racial integration in public schools, (5) government vs. private coverage of medical and hospital expenses, (6) immediate withdrawal from Vietnam vs. efforts to achieve complete military victory, (7) total vs. no government action to battle inflation, (8) protection of the legal rights of those accused of committing crimes, and (9) whether the federal government should make every effort to improve the social and economic position of blacks and other minority groups. On each seven-point scale for the respondent's self-placement on an issue, the score of "1" represented a position favorable to governmental action, progressive taxation, legalization of marijuana, busing, withdrawal from Vietnam, or protection of

the rights of the accused. A score of "7" indicated a position at the opposite extreme. By combining respondents' self-placements on each of the nine issues, we have a composite seven-point index of each voter's issue orientations.

In employing 1972 survey data to examine the relative impact of candidate images we use several measures of cognitive, affective, and conative aspects of images. For the cognitive aspect, i.e., belief about each candidate regardless of how one feels about or tends to react toward him, there are two sets of measures. First, in addition to being asked to identify what kind of liberal or conservative each respondent considered himself, the respondent--using again a seven-point scale--designated what he believed to be the respective ideological leanings of Richard Nixon and of George McGovern. Second, on each of the nine issues listed above, each respondent designated where he believed each candidate's stand was on the same seven-point scale he used to declare his own position. As in the case of respondents' self-placement on issues, we use a composite index to summarize their perceptions of each candidate's stands across all nine selected issues.

The affective aspect of candidate images consists of two measures of voters' feelings about contenders and likes and dislikes. First, we use the respondents' ratings of Nixon and McGovern on a "feeling thermometer" that taps the coolness or warmth of feelings on a scale of 0-100 degrees. Second, using the responses of people to an open-end standard SRC/CPS question asking for likes and dislikes about each candidate, we have a candidate index, a 13-point scale on which the higher the score the more pro-McGovern the respondent's affect. The index was calculated by taking, for each respondent, the number of things he/she liked or disliked about each candidate (up to three of each were coded in the survey); pro-Republican plus anti-Democratic responses were subtracted from pro-Democratic plus anti-Republican responses and the range of scores (-6 to +6) converted to a thirteen-point scale.

The conative aspect of images is more complex and warrants detailed consideration. Three decades ago Smith (1947) differentiated between three aspects of attitudes, or images--cognitions, affects, and conations. He stressed that a person's beliefs and feelings about any political object have much to do with what he proposes to have done about it. What one wants done Smith treated as conative and labeled "policy orientation." Following Smith and others, Harding, Proshansky, Kutner, and Chein (1969) examined the conative aspect of prejudice but broadened the notion to include people's "action orientations" as well as their ideas about "what should be done." They reported various studies measuring conation as the social distance perceived by people between themselves and members of various ethnic groups. This idea of distance, or proximity, we incorporate into our measurements as the conative aspect of images. Thus, with respect to a political candidate, the conative aspect is more than the voter's policy orientation in the sense of his intention to vote for or against the candidate. It refers also to how close a voter perceives himself to the political object, i.e., his proximity along specific dimensions (or with reference to specific matters) of sufficient salience to trigger for him an idea of what he proposes to do about the object. The 1972 survey provides two kinds of data to construct measures of proximity. First, respondents placed both themselves and each candidate on a seven-point ideological scale; it is possible to compare the absolute difference between the respondent's self-location and where he places a candidate and thereby derive a scale along which the higher a person's score, the more distant he considers his ideological position from that of Nixon or McGovern. Similarly, respondents rated themselves and each candidate on nine salient policy issues; we calculate an issue proximity index for each candidate indicating, as the

score increases, a growing distance between the respondent's issue stand and his perception of the candidate's stand.

TABLE 1

Summary Measures of Self, Candidate, and Party Images: 1972

---

Respondents' Self-images:

1. Partisan self-images (PSI)
2. Ideological self-images (ISI)
3. Issue orientations, or self-placement on issues (SPI)

Candidate Images:

Cognition:

4. Respondent's perception of Nixon's ideological position (NLC)
5. Respondent's perception of McGovern's ideological position (MLC)
6. Respondent's perception of Nixon's issue stands (NI)
7. Respondent's perception of McGovern's issue stands (MI)

Affect:

8. Respondent's rating of Nixon on feeling thermometer (NT)
9. Respondent's rating of McGovern on feeling thermometer (MT)
10. Index of respondent's likes and dislikes of the candidates, i.e., candidate index (CI)

Conation:

11. Respondent's proximity to Nixon's perceived ideological position (NLCP)
12. Respondent's proximity to McGovern's perceived ideological position (MLCP)
13. Respondent's proximity to Nixon's perceived issue stands (NIP)
14. Respondent's proximity to McGovern's perceived issue stands (MIP)

Party Images:

Cognition:

15. Respondent's perception of Republican party's ideological position (RLC)
16. Respondent's perception of Democratic party's ideological position (DLC)
17. Respondent's perception of Republican party's issue positions (RI)
18. Respondent's perception of Democratic party's issue positions (DI)

Affect:

19. Index of respondent's likes and dislikes of the political parties, i.e., party index (PI)

Conation:

20. Respondent's proximity to Republican party's perceived ideological position (RLCP)
  21. Respondent's proximity to Democratic party's perceived ideological position (DLCP)
  22. Respondent's proximity to Republican party's perceived issue positions (RIP)
  23. Respondent's proximity to Democratic party's perceived issue positions (DIP)
- 

In measuring images of the Republican and Democratic political parties we employed procedures parallel to those for measuring candidate images. The cognitive aspect of each party's image consists of the respondent's (1) perception of each party's position on the seven-point liberal/conservative scale and (2) perception of each party's stand on each of the nine policy issues (again combined into composite scales for each party). The affective aspects

of party images we tap with a party index derived from respondent's stated likes and dislikes about each party. And, we employ ideological and issue proximity measures similar to those used for candidates to measure the conative aspect by relying upon respondents' placements of each party on the liberal/conservative scale and on the nine policy issue scales. Table 1 summarizes the 23 measures and the labels used in reporting findings.

An initial question is whether any of the 23 variables relate to the vote. One way of answering this question is to correlate each image measure with respondents' voting behavior (using dummy variables of "0" for a Nixon vote and "1" for a McGovern vote). Space limitations make it impossible to present the full correlation matrix of each image variable with the vote and with one another. However, of the simple correlation coefficients between each of the image variables and the vote all but two proved statistically significant. Those two were respondents' perceptions of Nixon's issue stands ( $r=.05$ ) and of the issue stands of the Republican party ( $r=.08$ ). Of the significant coefficients, only a few are notably high. These include the ratings respondents gave to each candidate on the feeling thermometer ( $-.64$  and  $.64$  for Nixon and McGovern respectively), the respondents' proximities to each candidate's perceived ideological position ( $.56$  and  $-.50$  for Nixon and McGovern), the respondents' partisan ( $-.50$ ) and ideological ( $-.47$ ) self-images, and the respondents' proximities to what they perceive as the Republican party's ideological position ( $.46$ ).

In sum, simple correlations indicate that almost all of the 23 image measures relate to the vote and that selected political self-images, candidate images, and party images are especially noteworthy. Yet, simple correlations will not reveal which of these components are uppermost in the voters' decision-making processes. Nor does such analysis suggest the relative importance of the cognitive, affective, and conative dimensions of voters' images in helping them to make up their minds. To get at these matters factor analysis is a particularly useful tool.

Table 2 displays the results of a varimax analysis explaining 50% of total variance among all variables excluding the vote. The four-factor solution vividly refutes the original three-component model of voters' self, candidate, and party images. Factor I consists of nine variables, each pertaining to issues. Respondents' self-locations on issues, their perceptions of both candidate and party issue stands, and their proximities to candidates and parties on issues--all load on this factor and it thus exhausts all measures that in any way deal with issues. The second factor has ten variables and is bipolar, with five measures at the positive end and five at the negative. It is the most difficult factor to interpret for two of the items measure voters' self-images (partisan and ideological), three measure the affect aspect of candidate images (the two thermometers and the candidate index), the party index relates to affect toward parties, and the remaining variables are the respondents' proximities to the candidates' and parties' perceived ideological positions. If anything, this factor is a candidate-party factor. The remaining two factors offer no such interpretative problems. Factor III is a conservative factor consisting of respondent perceptions of the ideological positions of Nixon and the Republicans; the fourth factor is a liberal factor made up of perceived ideologies of McGovern and the Democrats. A second factor analysis with the vote added to the other 23 variables rounds out the picture of the principal components of the 1972 vote decision. With the exception of the vote, which enters as the highest loading variable on the second factor, the same four factors emerge as in the first analysis.

TABLE 2

Factor Loadings of Measures of Image Variables Excluding Vote Behavior

Variables	Factor Loadings				$\underline{h^2}$
	I	II	III	IV	
RI	.88	.02	.04	-.02	.78
NI	.87	.02	.04	-.02	.77
DI	.83	.02	.04	-.02	.69
DIP	.79	-.30	.01	-.05	.72
RIP	.74	.43	.07	.02	.75
MIP	.72	-.37	.00	-.06	.65
MI	.71	.24	.06	.00	.57
NIP	.64	.48	.07	.03	.77
SPI	.59	-.43	-.01	-.06	.53
MT	-.02	.63	.04	.07	.40
CI	-.02	.62	.04	.07	.39
NLCP	.00	.62	.32	-.08	.42
MLCP	.03	-.61	.14	-.17	.49
NT	.02	-.60	-.04	-.06	.36
RLCP	.00	.58	.30	-.08	.44
PI	-.02	.56	.04	.06	.32
PSI	.02	-.55	-.04	-.06	.30
ISI	.02	-.55	-.04	-.06	.30
NLC	.04	.13	.67	-.17	.49
RLC	.04	.12	.64	-.16	.45
MLC	-.03	.22	-.21	.43	.27
DLC	-.03	.22	-.21	.42	.27

The results of factor analyses warrant two assertions. First, the preconceived categories of the components of electoral decisions summarized in Table 1 are not distinct dimensions in the factor analyses. Nor are the electoral components derived from earlier voting studies (especially party, candidate, and issue orientation) clear-cut. To be sure, issue orientations stand out. However party and candidate orientations meld, and conservative and liberal orientations join the list as identified dimensions of the voting decision. Second, the vote appears on the candidate-party factor. Combining most strongly with it are three affective measures--i.e., the ratings of McGovern and Nixon on the candidate thermometers and the candidate index. This suggests a strong relationship between affective orientations toward the candidates and the vote in 1972, perhaps even that these affects were the overriding components of the electoral decision. Further to unravel some of the intricacies in the complex relationship between these electoral components and voting, we turn to the third and final stage of analysis, multivariate regression.

The results of a stepwise regression analysis using vote behavior as a dependent variable and the 23 image measures as independent variables appears in Table 3.

TABLE 3

## Regression Analysis of Image Variables Explaining 1972 Presidential Vote

Variable	Simple R	Multiple R	R <sup>2</sup>	Change	Beta
NT	-.64	.642	41.3%	.413	-.29
MT	.64	.768	58.9	.176	.30
NLCP	.56	.786	61.8	.029	.15
MLCP	-.50	.794	63.1	.013	-.10
PSI	-.50	.799	63.8	.007	-.11
CI	.48	.801	64.1	.003	.06
NIP	.36	.802	64.4	.003	.09
DIP	-.17	.804	64.6	.002	-.06
DLC	.18	.805	64.8	.002	.05
PI	.36	.806	64.9	.001	-.05
SPI	-.28	.807	65.0	.001	-.05
NLC	.14	.807	65.1	.001	.03
RLC	.12	.807	65.2	.001	-.02
MLC	.17	.807	65.2	---	-.03
NI	.05	.807	65.2	---	.02
ISI	-.47	.808	65.2	---	.02
MI	.20	.808	65.2	---	.03
DI	.11	.808	65.2	---	-.03
RIP	.32	.808	65.2	---	.02
RLCP	.46	.808	65.2	---	-.01
RI	.09	.898	65.2	---	.01
MIP	-.25	.808	65.2	---	-.01
DLCP	-.36	.808	65.2	---	.00

The Beta coefficients indicate that numerous variables had some independent effect upon the vote, but the relative importance of most (such as the respondent's proximity to the Democratic party's perceived ideological position) was virtually nonexistent. The key component of the vote was the affective evaluation of the candidates on the feeling thermometers. The affective orientation to Nixon explained 41% of the variance in the vote and affect toward McGovern added another 18%; the Beta coefficients of both are strong: -.29 for Nixon and .30 for McGovern. (Since the vote was coded as "0" for Nixon and "1" for McGovern, the negative Beta for Nixon indicates the negative correlation between his thermometer ratings and the vote.) Following affective orientations toward the candidates in importance are two variables originally conceived of as conative aspects of the candidate images and which factor analyses indicated as belonging to a candidate-party factor--that is, respondents' proximities to each of the candidate's perceived ideological predilections. Finally, the only other variable explaining an additional one per cent variance is the voter's partisan self-image (Beta = -.11). Thus, in a regression solution in which all independent variables explain a total of slightly less than two-thirds of the variance in the 1972 vote, five variables provide 64%. All five of these were on the candidate-party factor and four pertain directly to preconceived categories of affective and conative aspects of candidate images.

Composite indices of the perceived issue stands of the candidates and of the parties probably obscure some of the differences that voters detect between



contestants on specific issues. This may account for the fact that issues did not appear in this analysis as relatively important explanations of the 1972 vote. Another regression analysis helped test for this possibility. This regression maintained the separation of five key issues (thus, entering respondents' self-locations, perceived candidate stands, and proximities to candidates on each issue as separate variables in the analysis). Only the voters' proximities to the perceived Republican position on the Vietnam issue (addition of four percent variance and Beta .20) and proximities to McGovern on Vietnam (one percent variance and Beta = -.16) had relatively important effects. Thus, the farther respondents deemed themselves from perceived Republican policy on Vietnam and the closer to McGovern's perceived stand, the more likely they voted for McGovern. Again the thermometer ratings emerged as the two most important variables; in a regression explaining 72% of the variance in vote the feeling ratings explained 59%. Since regression analysis permits a probe of only the direct relationships between selected variables and the vote, we remain uncertain as to what indirect effects such influences as partisan and ideological self-images have as they condition voters' perceptions of candidates. The moderate correlations between these variables and affective orientations toward the candidates lead us to suspect such an indirect, conditioning effect. Moreover, other analyses comparing the direct and indirect effects of partisan self-images on the 1972 vote suggest considerable indirect effects (Miller, 1972).

In sum, despite the restrictions inherent in available measures and techniques there are at least partial answers to the questions posed at the outset of this presentation. First, most of the 23 image measures exhibit a direct relationship to the vote. Second, there are discrepancies between earlier studies' conceptions of the components of the vote as being the voter's self-image, candidate images, and party images and the principal components derived from reported factor analyses. Instead, the components appear to be issue, candidate-party, and liberal-conservative orientations. Yet, step-wise regression indicates that candidate image concerns on the candidate-party factor, especially in their affective and conative aspects, are the most important explanations of the 1972 vote. Unanswered, of course, are such questions as the relative magnitude of indirect effects on candidate images and voting behavior of other influences such as party identification.

#### REFERENCES

- Boulding, K. E. The image. Ann Arbor: University of Michigan Press, 1956.
- Boyd, R. W. Presidential elections: an explanation of voting defection. American Political Science Review, 1969, 63, 498-514.
- Butler, D. & Stokes, D. Political change in Britain. New York: St. Martin's Press, 1969.
- Campbell, A., Gurin, G. & Miller, W. E. The voter decides. New York: Harper, Row and Co., 1954.
- Campbell, A., Converse, P., Miller, W. & Stokes, D. The American voter. New York: John Wiley & Sons, 1960.
- Campbell, A., Converse, P., Miller, W. & Stokes, D. Elections and the political order. New York: John Wiley & Sons, 1966.

- Downing, J. What is a brand image? Advertising Quarterly, 1964-65, 1, 13-19.
- Harding, J., Proshansky, H., Kutner, B. & Chein, I. Prejudice and ethnic relations. In Lindzey, G. & Arenson, E. (Eds.), The handbook of social psychology. Vol. 5. Reading, Mass.: Addison-Wesley Publishing Co., 1959.
- Jensen, R. American election campaigns: a theoretical and historical typology. Paper presented at the meeting of the Midwest Political Science Association, Bloomington, Indiana, 1968.
- Kelley, S. & Mirer, T. The simple act of voting. American Political Science Review, 1974, 68, 572-91.
- Kovenock, D., Beardsley, P. & Prothro, J. Status, party, ideology, issues, and candidate choice: A preliminary, theory-relevant analysis of the 1968 American presidential election. Paper presented at the meeting of the International Political Science Association, Munich, 1970.
- McGinniss, J. The selling of the president 1968. New York: Trident Press, 1969.
- Miller, A. H., Miller, W. E., Raine, A. S. & Brown, T. A. A majority party in disarray: Policy polarization in the 1972 election. Paper presented at the meeting of the American Political Science Association, New Orleans, 1973.
- Miller, R. E. & Jackson, J. S. Split ticket voting: An empirical approach derived from psychological field theory. Paper presented at the Southern Political Science Association, Atlanta, 1973.
- Minow, N. M., Martin, J. B. & Mitchell, L. M. Presidential television. New York: Basic Books, Inc., 1973.
- Natchez, P. B. Images of voting: The social psychologists. Public Policy, 1970, 18, 553-87.
- Natchez, P. B. & Bupp, I. C. Candidates, issues, and voters. Public Policy, 1968, 17, 409-37.
- Smith, M. B. The personal setting of public opinion. Public Opinion Quarterly, 1947, 11, 507-23.
- Wyckoff, G. The image candidates. New York: The Macmillan Co., 1968.



ATTRIBUTION THEORY AND VOTER CHOICE<sup>1</sup>

Richard F. Yalch  
University of Washington

Voters frequently possess little campaign information but still satisfy their citizenship obligation by selecting candidates for various offices. Self-attribution concepts have been suggested as the appropriate internal information processing mechanism in situations in which attitudes have low saliency and are products of behavior rather than its antecedent. The assertion that voter behavior and attitudes reflect self-attributions was tested in a field experiment conducted during a local election. False feedback after a pre-election interview was hypothesized to be internalized as the subject's attitude toward voting and to influence actual turnout. The findings support the predictions and suggest campaign techniques to influence voter choice.

Theorists have tended to adopt one of two approaches when conceptualizing the voter decision process. The "rational" model portrays the voter as a recipient of a stream of communications describing the candidates' positions on the issues. He contrasts this with his own position and on the basis of the matching of desires and expectations decides if and for whom to vote. Perhaps the best known advocate of this approach is Downs (1957) in his influential Economic Theory of Democracy. On the other hand, the "emotional" voter model assumes that the voter is exposed more to nonissue information from which he develops a perception of the candidates' images. His ultimate decision is based on his affection for any of these images. Nimmo (1970) provides a thorough discussion of the art of image politics.

This characterization probably overemphasizes the differences in the approaches. Clearly, image can be defined to include the candidate's position on important issues and the political economist could include the candidate's personality as one of his attributes. Moreover, there is the common assumption that the voter is exposed to a considerable amount of campaign information, pays attention to it, understands it, and remembers enough of it so that it influences his ultimate choice. Although this assumption is reasonable for presidential and major state elections, there are reasons to doubt that it is appropriate for local elections. Anyone who has conducted interviews prior to a minor election or engaged in door-to-door canvassing is quickly impressed by the few voters knowing more than a candidate's name and party identification and the many not able to recall even that much.

Other evidence also suggests that local elections should be considered as different behavioral situations than the more often researched and analyzed national elections. Books on campaigning are quick to cite instances in which a candidate won an election because the voters had mistaken his name for that of a prominent person (see Nimmo, 1970, pp. 11-12). Realizing the limited information guiding the voters, candidates worry about their ballot position much the way brand managers are critically concerned about shelf space. In primary elections, the voters do not have the help of party identification to guide them and poor ballot positions have been known to cause the defeat of competent incumbents. To further support the contention that issues are not the decisive factor in local elections, one can cite Stokes and Miller's (1966)

conclusion that only about seven per cent of a national sample had "any discernible issue content" in the reasons they gave for voting Democratic or Republican in the 1958 congressional elections. Therefore, it is likely that issues are even less important in a city council race. A campaign manager trying to elect a candidate by promoting his positions or image would have an extremely difficult time convincing a majority of the voters. Consequently, most local campaigns concentrate on building name identification in the hope that the "mere exposure" hypothesis is valid.<sup>2</sup>

In a frequently cited article, Krugman (1965) distinguishes between the possible effects of a mass communication campaign on the basis of the target behavior. If it is a high involvement task, such as the selection of a president, the campaign's effectiveness can be determined by its immediate effect on the audience's attitudes which will determine behavior. This is the appropriate situation for applying a rational, emotional or combined voter decision model as the vehicle to identify the critical determinants of choice. However, for less involving behaviors, which the typical minor election vote decision is, messages may not have a noticeable effect on attitudes until the point of purchase or in usage. Krugman (1965) was not entirely certain how this occurred and, in fact, concluded his article by pointing out the need for a low involvement model to explain the relationship between attitudes and behavior in these situations.

One possible framework for analyzing low involvement behavior is attribution theory since it describes how individuals assign attitudes to behaviors when they are not sure of the exact causes (Kelley, 1967). Bem (1972) has extended this theory to explain the self-assignment of an attitude after a behavior. Specifically, the first proposition of his self-perception theory states that:

Individuals come to "know" their own attitudes, emotions, and other internal states partially by inferring them from observations of their own overt behavior and/or the circumstances in which this behavior occurs.

This proposition has received a compelling amount of support from research on many diverse behaviors and situations. For example, Lepper (1973) demonstrated that children's perception of their own honesty was a function of their appearing to behave honestly in a prior task; Ross, Insko and Ross (1971) found that subjects used the physical evidence of an initial questionnaire response as a guide in a second response despite the fact that the evidence had been altered so that it reflected an opposite attitude; McArthur, Kiesler and Cook (1969) used a false personality feedback to induce greater volunteering for a post-experimental task; and Schachter and Singer (1962) reported that subjects determined their emotional state after being drug-aroused by observing the state of a confederate in the same situation.

Although this research is impressive in its robustness and consistency, it has been rarely utilized in consumer behavior research. This may only reflect the general time lag in the "borrowing" process that has occurred between the behavioral sciences. However, it also might reflect a justified caution about adopting theories that have been primarily tested in laboratory settings, using student samples, self-report dependent measures, and inconsequential behaviors. One might rightfully question the relevance of the theory. Can it be used to alter "meaningful" behavior in a natural environment with a reasonable expenditure of resources?

This research study had several objectives. First, it attempted to test self-perception theory in a natural setting and to demonstrate that it suggests ways

to alter both attitudes and behavior. Inducing a change in both variables is a nontrivial task given the evidence that an attitude change does not lead to a behavior change (Festinger, 1964). Second, it explored the perplexing finding that attitudes toward the candidates in an election do not account for much of the variance in the decision to vote or abstain in an election (Yalch, in press). This involved determining whether an individual's perception of himself as a likely or unlikely voter was a major determinant of voting behavior. Lastly, a possible strategy for increasing voter participation in the electoral process was to be evaluated.

#### Procedure

The experiment was conducted during two special aldermanic elections in Chicago during the summer of 1973. The preliminary election featured five candidates and was used to pilot test the induction procedure. The main study occurred during the run-off election as part of a pre-election survey. The 190 subjects were selected from official lists of registered voters and contacted door-to-door. After the subjects had completed the survey, the interviewers pretended to score the responses on a bogus voter profile sheet. This was given to the respondents accompanied by a brief verbal explanation of its interpretation. The subjects' voting behavior on election day was assessed from party records and their self-perception as frequent or infrequent voters was determined in a post-election interview.

#### Treatments

There were two treatment groups and one control group. The 77 subjects randomly assigned to the above average treatment received the voter profile sheet illustrated in Figure 1. Their responses were supposedly represented by the red line drawn in after the interview. The position of the line to the right of the center or normal line was explained as indicating that their responses reflected a greater interest and concern about the election and politics than the average citizen. In addition, it was pointed out that this suggested that they would be more likely to vote in an election. The profile was left with them to further reinforce the information that they were above average citizens.

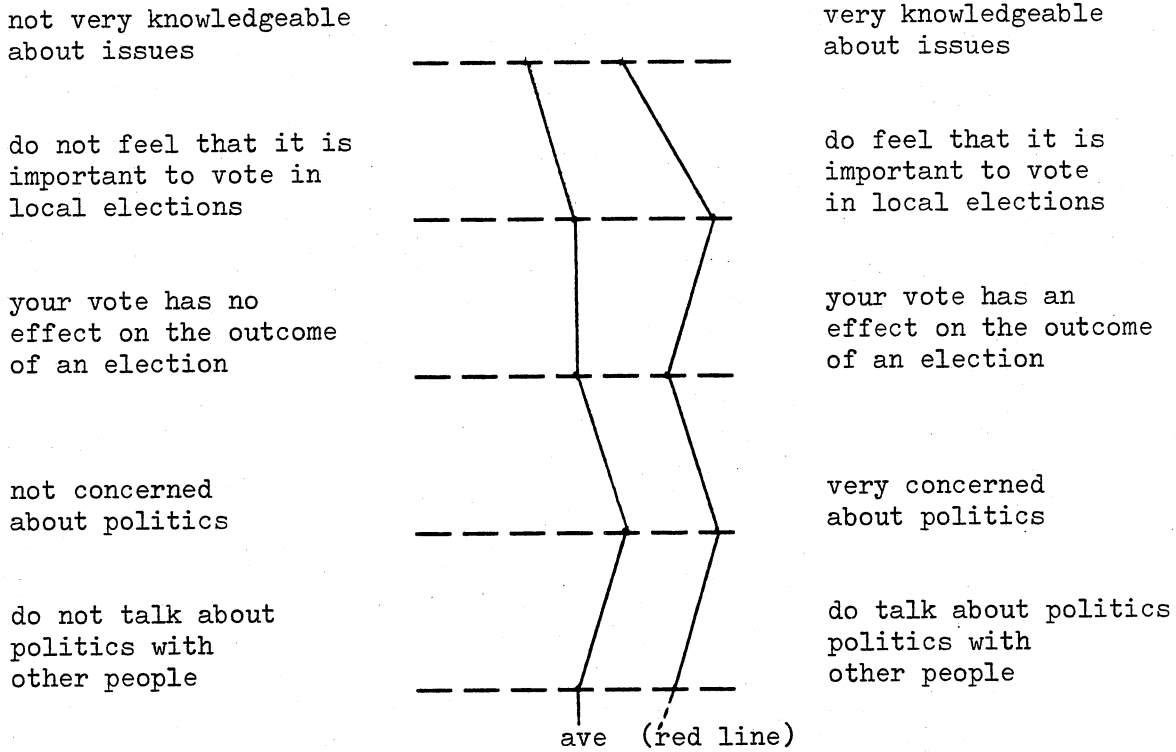
The average group consisted of 85 subjects receiving the profile sheet illustrated in Figure 2. It was identical to the other profile except that their response line was drawn near the center of the scale. This was interpreted to them as evidence that they were about the same in their political orientation and interest as the average person in the community. They were not told that they would be very likely to vote in the election, but they were also allowed to keep the profile sheet.

The 28 persons in the control group were interviewed door-to-door, answering the same questions as the treatment groups, but were not told anything about their disposition toward politics. This group's behavior enables one to estimate the absolute effect of being given feedback after an interview

#### Dependent measures

The treatment effects were assessed in two ways. First, the voting behavior of each subject was determined from records maintained by a campaign organization. The ability to have an unobtrusive dependent measure was a major reason why voter turnout rather than candidate choice was selected as the experimental behavior. This is advantageous in minimizing threats to the construct validity of the experiment. The second measure was the subjects' self-perception as a

Voter Profile for Chicago



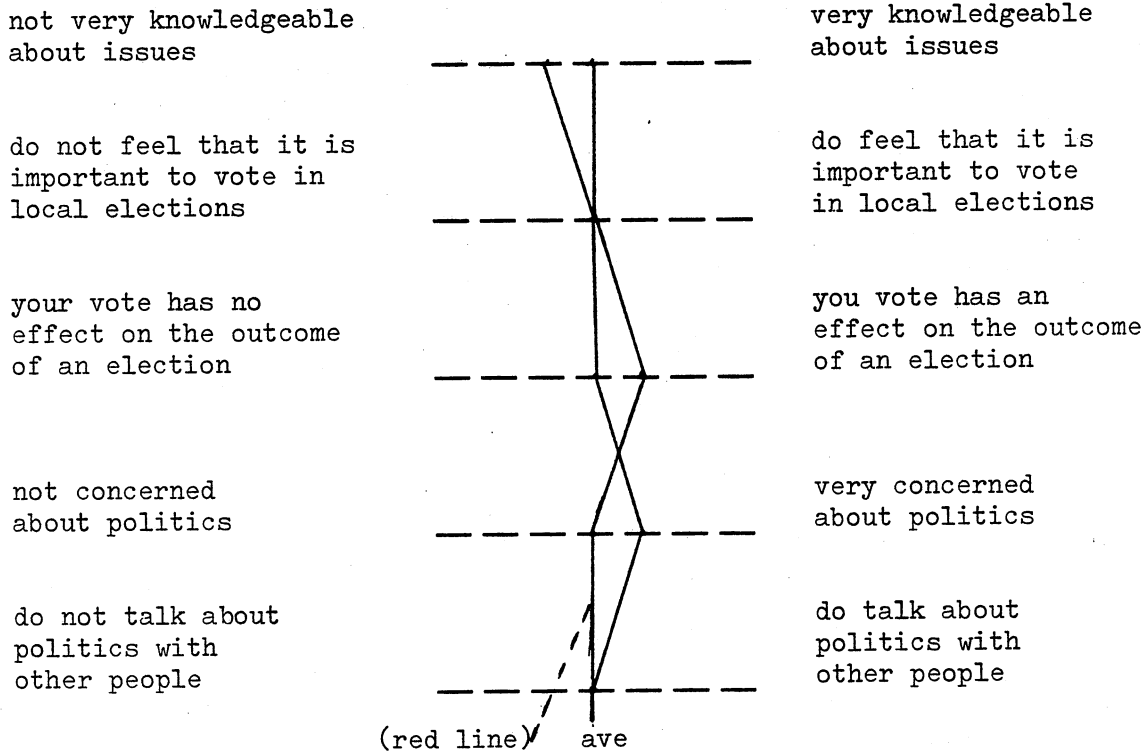
The profile of the average citizen in your community is indicated by the vertical black line connecting the points. It has been developed from telephone interviews conducted in May as a Northwestern University project.

Your responses are indicated by the red line. If this line is to the left of the black line at any point, this shows that your answers are most like those people least likely to vote. If the red line is to the right of the black line, this shows that your answers are most like those people most likely to vote.

Thank you for your cooperation.

Figure 1. Above average treatment instrument.

Voter Profile for Chicago



The profile of the average citizen in your community is indicated by the vertical black line connecting the points. It has been developed from telephone interviews conducted in May as a Northwestern University project.

Your responses are indicated by the red line. If this line is to the left of the black line at any point, this shows that your answers are most like those people least likely to vote. If the red line is to the right of the black line, this shows that your answers are most like those people most likely to vote.

Thank you for your cooperation.

Figure 2. Average treatment instrument.



frequent or infrequent voter. This was determined in a telephone survey conducted within weeks after the election. Several general questions were asked in addition to the self-perception question to minimize the possibility that the subjects would think that this was a test of their ability to recall the profile information.

### Hypotheses

1. Subjects given above average feedback will be more likely to report themselves as above average in the post-election interview than those given average feedback.
2. Subjects given above average feedback will be more likely to vote in the election than those given average feedback.

These two hypotheses were generated from self-perception theory's proposition that an individual's attitudes and behavioral dispositions reflect his past behavior. By altering his interpretation of this experience, it should be possible to alter his beliefs about himself. This in turn should influence his future behavior. No hypotheses were prespecified for the comparisons between the control group and the treatment groups because there was no information available as to how the residents of this community perceived themselves. Since this district usually has one of the highest turnout rates in the city, it was likely that many voters would naturally perceive themselves as above average. A tentative prediction was that the control group's attitudes and behavior would fall somewhere in-between the two treatment groups's.

### Results

Before the treatment effects were analyzed, a test of subject randomization was made. This involved comparing the preliminary election voting behavior of the groups (assessed from party records), and their intention to vote prior to receiving the feedback (assessed in the interview). There was a slight but statistically insignificant higher voter turnout in the above average and control groups compared to the average group. However, since there was only a weak association between voting in the preliminary election and voting in the run-off, this was not judged to be problematic. A more critical test was the difference in the voting intentions of the groups since this had a high correlation with subsequent voting. There was virtually no difference between the groups and it was concluded that the randomization had been successful.

Table 1 presents the post-election self-perceptions of the subjects. The difference between the above average and average feedback groups is statistically significant ( $p=.016$ ), using a test for the difference in proportions (Ferguson, 1973, pp. 160-162). The average group also had a significantly lower probability of saying that they were above average voters than the control group ( $p=.035$ ). Apparently, the control group did have a high self-perception of themselves as politically active citizens as their responses are equivalent to those from the above average treatment group.<sup>3</sup>

The behavioral effects of the treatments are presented in Table 2. The second hypothesis was supported as the above average group had a significantly higher turnout than the average feedback group ( $p=.041$ ). All comparisons with the control group were insignificant, partially as a consequence of the small size of the group. The finding that the average feedback group had a much lower self-perception and likelihood of voting relative to the control group suggests the possibility that telling persons that they are only average might undermine their intrinsic interest because one is inclined to think of

TABLE 1  
Self-Perception by Treatment<sup>4</sup>

	equal or less likely	more likely	%more likely
above average	16	48	75.0
average	28	36	56.2
control	5	18	78.3

himself as above average. Being told that this view is incorrect may have caused many in the average feedback group to question whether they really cared that much about voting in the election.

TABLE 2  
Voting Behavior by Treatment

	no	yes	% voting
above average	12	65	84.4
average	23	62	72.9
control	5	22	81.5

### Discussion

The experimental findings demonstrate the validity of self-perception theory and its applicability to local election voting behavior. There was clear evidence that voters used their past behavior and the interviewers' interpretation of it to determine their disposition toward voting in the election. A local election is a low involvement activity and persons usually have uncertain and unstable attitudes. Therefore, attitudes anchored to a prior behavior should have a more significant role in determining an individual's course of action. However, the behavior-attitude relationship appears to be more ambiguous to the individual than previously thought. The external information provided by the interviewer caused many persons who thought of themselves as only average to perceive themselves as above average, and conversely, many above average persons began to consider themselves as only average. The previous experiences which lead to this original self-perception were virtually negated by a single experience and the false information about it. The post-behavior attitude reassessment process represents a fruitful area for future consumer behavior research and a potentially very powerful method to influence behavior.

The study also provides further evidence for the belief that the decision to vote in an election is based on factors in addition to one's attitudes toward the candidates in the election. An individual's perception of himself as the type of person who votes frequently seems to be one of these factors. How this self-

perception develops is not yet clear, but childhood socialization and initial voting experiences probably play an important role. This finding helps to explain why many citizens continue to vote even though their candidates always lose. The act of voting in the absence of a tangible reward probably causes the individual to assume that he has a very strong citizenship attitude.

Since behavior appears to be a major determinant of one's attitudes and future behavior, political campaigns might benefit from a reallocation of resources from "information" dissemination to techniques designed to induce voters to make an early behavioral commitment to the candidate. The recent congressional elections featured a promotional strategy labelled "Dollars for Democrats". The ostensible purpose of the activity was to raise money for the campaigns, but it probably had an additional benefit in having voters make a small commitment to the Democratic party. On election day, a recall of this commitment could be influential in determining how one votes. Freedman and Fraser (1966) have demonstrated the importance of yielding to a small request on the likelihood of yielding to larger requests at a future date. The methods for inducing a favorable attitude without resorting to a flood of campaign propaganda are many and should be the object of future research and practice.

#### FOOTNOTES

1. The author wishes to thank Professor H. Woods Bowman, University of Illinois at Circle Campus and Northwestern University doctoral candidates Alice Tybout and Richard Bagozzi for assistance in the data collection.
2. The mere exposure hypothesis states that the repeated presentation of a neutral stimuli is a sufficient condition for increased affection for it. Thus, just seeing a candidate's name frequently should cause one to gradually favor him (Zajonc, 1968).
3. A comparison of the voting behavior of the control group with another sample interviewed by telephone suggests that this group may have contained more politically involved citizens than the other groups. The above average feedback, therefore, probably did increase the group's self-perception more than is indicated by comparisons with the control group.
4. Several subjects could not be recontacted after the election so the numbers in this table are smaller than the number participating in the experiment.

#### REFERENCES

- Bem, D. J. Self-perception theory. In L. Berkowitz (Ed.), Advances in experimental social psychology. New York: Academic Press, 1972.
- Downs, A. An economic theory of democracy. New York: Harper and Row, 1957.
- Ferguson, G. A. Statistical analysis in psychology and education. New York: McGraw-Hill, 1973.
- Festinger, L. Behavioral support for public opinion change. Public Opinion Quarterly, 1964, 28, 404-417.
- Freedman, J. L. & Fraser, S. C. Compliance without pressure: the foot-in-the-door technique. Journal of Personality and Social Psychology, 1966, 4, 195-202.

- Kelley, H. H. Attribution theory in social psychology. In D. Levine (Ed.), Nebraska symposium on motivation. Lincoln: University of Nebraska Press, 1967.
- Krugman, H. E. The impact of television: learning without involvement. Public Opinion Quarterly, 1965, 29, 349-356.
- Lepper, M. R. Dissonance, self-perception, and honesty in children. Journal of Personality and Social Psychology, 1973, 28, 65-74.
- McArthur, L., Kiesler, C. A. & Cook, B. Acting on an attitude as a function of self-percept and inequity. Journal of Personality and Social Psychology, 1969, 12, 295-302.
- Nimmo, D. The political persuaders. Englewood Cliffs, New Jersey: Prentice-Hall, 1970.
- Ross, M., Insko, C. A. & Ross, H. Self-attribution of attitude. Journal of Personality and Social Psychology, 1971, 17, 292-297.
- Schachter, S. & Singer, J. E. Cognitive, social, and psychological determinants of emotional state. Psychological Review, 1962, 69, 379-399.
- Stokes, D. E. & Miller, W. E. Party government and the saliency of Congress. In A. Campbell, P. E. Converse, W. E. Miller & D. E. Stokes (Eds.), Elections and the political order. New York: John Wiley & Sons, 1966.
- Yalch, R. F. The prediction of voter turnout. In G. Zaltman & B. Sternthal (Eds.), Broadening the concept of consumer behavior, in press .
- Zajonc, R. Attitudinal effects of mere exposure. Journal of Personality and Social Psychology Monograph Supplement, 1968, 9, 1-27.



APPLICATIONS OF MARKETING CONCEPTS TO CANDIDATE MARKETING<sup>1</sup>

Avraham Shama<sup>2</sup>  
Baruch College/CUNY

Sellers, products, consumers, market segmentation, product image, brand loyalty, product development, product concept, product positioning, market research, and concept testing are basic marketing concepts which are applicable to political marketing. An examination of the applicability of these concepts to candidate marketing, and a detailed examination of consumer and voter behavior suggest the inclusion of political marketing within the boundaries of marketing theory.

Political marketing is the process by which political candidates and ideas are directed at the voters in order to satisfy their political needs and thus gain their support for the candidate and ideas in question. A cursory comparison between marketing of goods and services, and marketing of political candidates would readily point up at least one common concept: promotion, viz, the apparent and quite extensive use of media by the seller and the candidate for the purposes of informing, reminding, attitude-affecting, and sales-facilitating activities focused on target groups of buyers and voters, respectively. Possibly, such a comparison would also indicate that both marketing of goods and services, and marketing of political candidates utilize similar tools such as market research, and various statistical and computer techniques in studying the market. Although these points are essentially correct, they denote only a few of the similarities between marketing and political marketing.

A more serious comparison, however, will indicate that many more concepts and tools are shared by marketing of goods and services, and marketing of political candidates. Consider, for example, some well-known concepts of marketing: sellers and buyers, consumer behavior, market segmentation, image, brand loyalty, product concept, and product positioning, etc. They are all concepts of political marketing. Consider also some of the familiar tools which are used in marketing: market research, media, advertising, multiple regression, factor analysis, discriminant analysis, conjoint measurement, and multi-dimensional scaling, etc. They are all tools utilized in the marketing of political candidates (Kotler, 1975).

But perhaps the most powerful test for applying the concept of marketing in the area of political marketing, is by examination of the applicability of consumer behavior concepts to the area of voter behavior. The reason for this is because the consumer orientation of marketing has made consumer behavior concepts the focal points of marketing.

#### Marketing and Political Marketing

##### Similarities of Concepts

Common concept one: sellers, products, and buyers. Both marketing and polit-

ical marketing include three main elements: sellers, products, and buyers. Marketing is a process by which sellers offer the buyers products and services in return for something of value (usually money). The same process takes place in political marketing, whereby the candidates offer the voters products or ideas such as "economic prosperity," "safe society," etc. in return for their votes and support in the campaign period and thereafter. The fact that many economic products can be sold and bought often while buying the product that political candidates offer can be done only infrequently and at a fixed point in time and space does not invalidate this argument, but rather indicates differences in nature and use of political and economic products, very similar to the differences in nature and usage of products and services which are traditionally subsumed by marketing (e.g.: food items vs. durable goods, insurance, auctioned merchandise).

Common concept two: consumers. The core of both marketing and political marketing are the consumers. Without consumers, the marketer of economic goods and services does not have a market, and without voters the political marketer does not have a campaign. Because both marketers need consumers to survive, the concept of consumer behavior or voter behavior becomes a focal point of marketing and political marketing, respectively. The fact that in one case an individual is called "consumer" and in another "voter," is merely a semantic difference. In both cases the individual can be viewed as an organism receiving stimuli about the product and reaching predispositions to respond, and a final response state after going through an essentially similar decision making process. Accordingly, the principles of well known models of consumer behavior can certainly be applied to voter behavior, and vice versa. In fact, the similarities here are so strong, that consumer behavior literature and models perhaps unknowingly relate to concepts which were first developed in the literature of voter behavior, e.g.; selective exposure, selective perception, two-step flow of communication, etc.<sup>3</sup>

Common concept three: market segmentation and product mix. Both marketing and political marketing utilize the concepts of market segmentation and target groups to increase sales and votes, respectively. Market segmentation is the process by which consumers and potential consumers of the product are distinguished along one or more variables so as to create homogeneous groups, and select some of them as target groups in order to offer a satisfactory product mix, and achieve the company's goals (e.g.; profit, growth, market share, etc.) Variables along which product and candidate markets are segmented are almost identical: age, sex, income, occupation, family size, race, personality characteristics, life style, etc. Furthermore, product-specific variables such as previous product use and preferred product characteristics are often similarly used (e.g.: "how many times did the voter support the same program or candidate before?" "What does the voter like most about the candidate?" etc.) As target groups the product marketer and the political candidate select consumers and voter, respectively, and offer them satisfactory product mixes. The product mix, viz., the different mixes of product, promotion, price and place that are offered to different voter segments, is also similar to the idea of product mix of marketing.

Common concept four: product image. Both product marketing and candidate marketing have emphasized that consumer and voter behavior toward products and candidates is shaped by their images of the products and the candidates in question. In addition, it seems that they both have overpopularized the

image concept to a degree where it became merely an impression or a stereotype that consumers and voters have about the products and candidates, respectively.

Common concept five: brand loyalty. Measured by the degree of attachment to the brand (as indicated by repeated purchase or brand attitude), and related to such consumer's characteristics as age, income, race, personality, etc., and in turn simplifies the decision making process of the consumer, brand loyalty becomes equivalent to the concept of party loyalty of political marketing. Furthermore, the concepts of brand loyalty and party loyalty have been utilized as a baseline for promotion strategy for the product and the candidate. Accordingly, the first step of such promotion strategy is to distinguish between voters who are loyal to the party and swing voters (i.e.: brand loyalty and brand switching), and hence design a different promotion mix for each of the two main groups. (Kotler, 1971 and Campbell et. al., 1966).

Common concept six: product development. Both product marketing and political marketing place great importance on the series of integrated activities and research that take part in the process of developing a product that will satisfy the target consumers and voters, respectively. In the case of consumer products, product development is a process through which a consumer-satisfying parcel of ingredients, quality, brand, package, etc. is created. Similarly, the process of developing a product in the political market is one of creating a parcel encompassing a candidate, issues, party, etc., which will satisfy the target voters.

Common concept seven: product concept. Essentially a part of the product development process, product concept, viz., the central idea(s) which serves as the core of the product in the target group's mind, is shaped by marketing and political marketing. Thus, an economic product such as a car might be planned and developed to convey "economy" and "dependability," and be perceived as such by its target consumers, while a candidate might wish to convey "healthy economy" and "active foreign policy," and be perceived as such by his target voters.

Common concept eight: product positioning. Related to the above concepts of product development and product concept, the idea of product positioning, viz., the process by which the product is positioned vis-a-vis its competitors in the market, is clearly utilized by both marketing and political marketing. In both cases, the product's and the candidate's "location" in the perceptual map of consumers and the voters relative to the "location" of the competitors is to be determined, planned, and promoted so as to increase consumer and voter preference of the product and the candidate in question. In addition, products and candidates utilize the same research technique in determining and planning their positions in the market in relation to their competitors, namely multidimensional scaling.

#### Common Tools

Common tool one: market research. Both marketing of economic products and services, and the marketing of political candidates make frequent use of market research or public opinion polling for the purposes of measuring product performance, identifying potential consumers, and detecting and solving problems. In doing so, market research and opinion research use similar methods of data collection, e.g.: panels, interviews, questionnaires, etc., and data analyzing techniques, e.g.: correlations, regressions, factor



analysis, discriminant analysis, multidimensional scaling, etc.

Common tool two: concept testing. A technique used in the process of product development and product positioning, concept testing refers to the procedure which is designed to discover consumer reactions to different product concepts or ideas so as to help management choose a suitable product concept, develop and introduce it to the market to satisfy the target consumers. Although not to the same degree of sophistication, this procedure is used by both marketing and political marketing. Thus, similar to product concept-testing, the candidate concept-testing involves the following major steps: (1) identification of possible candidate concepts; (2) introduction of candidate concepts to the voters; (3) recording voter reactions to each concept (by rank order, attitude measurement, intentions to vote, etc.); (4) identification of causal or associative connections among voters characteristics (SES, behavioral, and political) and their reactions to different candidate concepts, and to various attributes of single candidate concepts (factor analysis to reduce candidate attributes space, and voter characteristics space, and multiple regressions) so as to evaluate the contribution of separate candidate attributes and voter characteristics in the overall preference or ranking of the concept; (5) choice of the most positively evaluated candidate concept or concepts, and finally (6) introduction and promotion of the chosen concept or concepts among voter groups in reference to the results in step four. Clearly, this also implies the possibility for market segmentation: different voter groups who evaluated a given candidate concept differently can be regarded as different market segments.

Common tool three; communication. The instrumental use of communication media for the purpose of promoting economic products and political candidates is another characteristic of both marketing and political marketing. Each of these utilizes media schedules, and media mix to effectively reach its target groups. Consistent with this last point, the fact that product marketing and candidate marketing use different media mixes or schedules should be regarded as an indication of the different nature of the products, and their target groups.

One clear conclusion that can be drawn so far is that marketing as it has been traditionally conceived to refer to economic products and services, and political marketing, which relates mainly to marketing of political candidates, have much in common: (1) basic concepts such as sellers, buyers, products, consumers, market segmentation, etc. that are the core of each of them, and (2) tools or techniques used in market research and opinion research, concept testing and use of media. On these grounds, the concept of marketing seems to be quite applicable to the area of political marketing.

However, a more critical examination of broadening the concept of marketing to include also political marketing is provided by examining the applicability of the most important concept of modern marketing, namely that of consumer behavior, to the area of voter behavior.

#### Consumer and Voter Behavior

Voter behavior has been studied much in the same manner as consumer behavior, namely as a decision making process to engage in a certain action (voting, purchasing), including processes which precede and follow

that act. Both the voter and the consumer are viewed as individuals receiving information, and possibly seeking out information, processing this information to reach predispositions to respond, and finally responding toward the product and the candidate in question. Consequently, the principles of well known models and frameworks of consumer behavior can be effectively applied to voter behavior and vice versa. Accordingly, in applying the general approach of consumer behavior models to voter behavior, one can point out the following components that are part of the decision process:

1. Stimulus input variables which originate from the candidate and his party and are targeted at the voters. Such input variables may relate to the candidate's experience in politics, his style of action as a political figure, his stands on issues, his party identification, etc.
2. Environmental influences on the voter. These relate to such factors as social class, peer group, and family influence on the voter, as well as the influence of the voter's own personality traits, and past experience with the candidate in question.
3. Processing stimulus and environmental information to reach voting predispositions. Such processing is subject to learning and selectively screening.
4. Output variables which relate to the decision how to vote, as well as to changes in perception of, and attitude toward, the candidate. One of the most powerful output variables is the voter party identification which, in a manner similar to brand loyalty, denotes an attachment to the party, and therefore also to its candidates.
5. Feedback processes.

Similarly, in applying voter behavior approaches to consumer behavior, one might follow the approaches of Lazarsfeld *et. al* (1948, 1954) and Campbell *et. al.* (1957, 1966) and postulate that consumer behavior is determined by socioeconomic status and psychological makeup. To be more detailed, one can follow Lane's S→O→R model (1965, P. 6), and describe the consumer decision making process as including the following three components (S, O, R):

1. The stimuli which are transmitted to the consumer from his social environment (e.g.: the community, media, family, ethnic group, social class, marketing channels) about the product or service in question.
2. The organism or the consumer receiving such stimuli, screening them through his perceptual and attitudinal screens in order to reach predispositions to respond toward the product or the service.
3. Responses such as purchasing, expressing an opinion about the attitudinal object, reading and listening to messages about the attitudinal object, etc.

In addition to these almost interchangeable models utilized by consumer and voter behavior scientists, the effectiveness of such models

in both areas is also very similar. Thus, consumer behavior models and voter behavior models can equally describe and to some degree explain consumer and voter behavior, but both fall short of predicting such behavior. Consequently, both consumer and voter behavior scientists seem to prefer the use of a middle-range theory approach in analyzing consumers and voters, respectively, rather than the reliance on fully blown theories in fields not yet ready enough for them. This takes the form of conducting studies which focus on the relationship between a fairly defined concept such as social class, family, peer group, etc., or a limited number of variables such as self-confidence, education, and dogmatism, and consumer behavior or voter behavior. A summary of such a middle-range theory approach and findings in the areas of consumer and voter behavior is presented in Table 1. An examination of this table shows that different concepts such as learning, perception, and social class help to explain consumer and voter behavior in similar ways. Thus, for example, learning concepts help to explain the processes by which consumers and voters develop brand and party loyalty, respectively; perception contributes to the explanation of consumer and voter imagery and its resultant influence on their behavior; and social class help to explain the purchase of certain products and voter party orientation. Furthermore, consumer and voter behavior are of interest to three conceptually similar groups: (1) marketers, parties, and candidates, (2) consumers, and voters, and (3) policymakers. However, it can be argued that because most of the concepts depicted in Table 1 are sociological or psychological, it is their nature to help explain behavior, be it the consumer, the voter, the employee, or the child, and therefore the argument that such concepts help to explain consumer and voter behavior is valueless. While this last point seems to have a high degree of face validity, it is nevertheless a fallacy. Firstly, it is argued that concepts such as learning and perception help to explain consumer and voter behavior in a similar manner. Evidently, when a concept helps to explain two seemingly different behaviors, then at least it is more uniquely connected to these behaviors than to other kinds of behavior. Secondly, as consumer behavior and voter behavior are a focus of interest of three conceptually similar groups (marketers, consumers and voters, and policymakers), it is reasonable to assume close bonds between these two roles of human behaviors. But more importantly, when the last two arguments are incorporated with the suggestion that consumer and voter behavior models and approaches are interchangeable, one cannot avoid the proposition that the structure of consumer behavior and the structure of voter behavior are very similar, and therefore they can be treated as essentially one structure.

TABLE 1

Concepts And Findings In The Areas of Consumer Behavior and Voter Behavior

Concept	FINDINGS	
	Consumer Behavior	Voter Behavior
Learning	Brand loyalty Advertising exposure	Party identification Advertising exposure

TABLE 1 (Cont'd.)

Concept	FINDINGS	
	Consumer Behavior	Voter Behavior
Perception	Purchasing products and shopping in stores whose images are consistent with self-image	Voting for candidates whose images are consistent with self-image
Attitudes	Purchasing behavior maintains cognitive consistency	Voting behavior maintains cognitive consistency
Personality	Personality traits are associated with purchase behavior	Personality traits are associated with voting behavior
Motivation	Products and services satisfy psychological needs	Candidates satisfy psychological needs
Social Class	Purchase of certain products (e.g.: clothing, home furnishing, leisure products)	Party identification (to some extent): Blue collar--Democratic, Businessmen and Professionals--Republican
Peer Group	Purchase of certain products and brands (e.g.: cars, cigarettes, drugs)	Vote casting, reinforcement of voting decisions consistent with group orientation
Opinion Leaders	Influence purchase decisions	Influence voting decisions
Family	Life cycle is associated with purchase of some products (e.g.: durable goods, education, health care)	Life cycle is associated with liberal-conservative orientations
Age	Older: Stronger brand loyalty	Older: strong party identification
Education	Higher: weaker brand loyalty	Higher: weaker party identification
Personal Selling (canvassing)	Increases sales	Increases voter turnout
T.V. Advertising	Influences consumers with low involvement with the product	Influences voters with low involvement in politics

TABLE 1 (Cont'd.)

Concept	FINDINGS	
	Consumer Behavior	Voter Behavior
Brand loyalty	Market segmentation strategy	Market segmentation strategy

## FOOTNOTES

1. The author would like to acknowledge the comments of Professor Philip Kotler on an earlier version of this paper.
2. Avraham Shama is an Assistant Professor of Marketing, Baruch College, 17 Lexington Avenue, N.Y., N.Y. 10010.
3. See, for example, the similarities in approach in the following political and consumer behavior literature: Paul F. Lazarsfeld, Bernard R. Berelson, and Hazel Goudet, The People's Choice, New York: Columbia University Press, 2nd ed., 1948; Bernard R. Berelson, Paul F. Lazarsfeld, and William McPhee, Voting: A Study of Opinion Formation in a Presidential Election, Chicago: University of Chicago Press, 1954; Angus Campbell, Philip E. Converse, Warren E. Miller, Donald E. Stokes, The Voter Decides, Evanston: Row, Peterson and Co., 1957; Francisco Nocosia, Consumer Decision Process, New York: Prentice-Hall, 1966; John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior, New York: John Wiley and Sons, 1968; James F. Engel, David T. Kollat, and Rogers D. Blackwell, Consumer Behavior, New York: Holt, Rinehart and Winston, Inc., 2nd ed., 1973.
4. For simplicity, Howard-Sheth model of buyer behavior is utilized. See: John A. Howard and Jagdish N. Sheth, The Theory of Buyer Behavior, New York: John Wiley and Sons, 1968, p. 30.

## REFERENCES

- Berelson, B. R., Lazarsfeld, P. F. & McPhee, W. Voting: A study of opinion formation in a presidential election. Chicago: University of Chicago Press, 1954.
- Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. The voter decides. Evanston: Row, Peterson and Co., 1957.
- Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. Elections and the political order. New York: John Wiley and Sons, Inc., 1966, pp. 9-12.
- Engel, J. F., Kollat, D. T., & Blackwell, R. D. Consumer behavior. New York: Holt, Rinehart and Winston, Inc., 2nd ed., 1973.
- Howard, J. A. & Sheth, J. N. The theory of buyer behavior. New York: John Wiley and Sons, 1968.
- Lane, R. E. Political life: Why and how people get involved in politics. New York: The Free Press, 1965.
- Lazarsfeld, P. F., Berelson, B. R. & Goudet, Hazel. The people's choice. (2nd ed.)

New York: Columbia University Press, 1948.

Kotler, P. Marketing management: Analysis, planning, and control. Englewood Cliffs, New Jersey: Prentice-Hall, 1974.

Kotler, P. Marketing for nonprofit organizations. Englewood Cliffs: Prentice-Hall, Inc., 1975, Chapter 19.

Nocosia, F. Consumer decision process. New York: Prentice-Hall, 1966.



RELATIONSHIPS BETWEEN CONSUMERS' SHOPPING AND LEISURE ACTIVITIES  
AND THEIR ATTITUDES TOWARD THE ENERGY CRISIS: A CROSS SECTIONAL STUDY<sup>1</sup>

Glenn S. Omura<sup>2</sup> and W. Wayne Talarzyk<sup>3</sup>  
The Ohio State University

Reported here is a descriptive analysis of a series of attitudinal questions relating the energy crisis to consumers' stated shopping patterns and leisure-time activities. The analysis focuses on consumer responses to key questions and those classification variables that best discriminate between those who stated their activities were affected or behavior changed and those who stated there were no effects or changes.

During the past twelve months a social issue known as the energy crisis has affected most Americans in many ways. This issue has provided an opportunity and, perhaps even more so, an obligation for consumer behavioralists to study the impact of a major change agent upon the attitudes, activities and behavioral patterns of large segments of consumers.

There are no easy answers to the multitude of changes and problems brought about by the energy shortage. In fact, even the questions that need to be explored are somewhat difficult to identify and isolate.

Therefore, while no single research project can bring all aspects of the energy crisis into perspective, hopefully such studies as the present collectively can provide a composite picture of what consumers think about the issue as well as how they have been affected by it. This paper is designed to fit into this composite framework by examining some statements related to shopping patterns, leisure activities and associated attitudinal questions.

#### Objectives

This paper reports on a portion of a larger data set on the energy crisis as overviewed in an earlier paper by the same authors (Talarzyk and Omura, 1974). Figure 1 illustrates the major components of the overall study.

It was reported in the earlier paper that one of the major objectives for the energy crisis research project was "to determine and evaluate what consumers think about the energy shortage and how they have been affected by it." The data to be examined here further serve to meet the above objective by focusing on two areas of Block C in Figure 1, "Leisure Time" and "Shopping Behavior," as they relate to "Attitudes Toward and Perceptions of the Energy Crisis" (Block B) and "Socio-Economic Variables" (Block D).

Basically the purpose here is to examine the relationships between consumers' attitudes toward the energy crisis and certain leisure-time activities and shopping patterns. It is also of interest to determine any differences between those consumers who said they were affected and those who



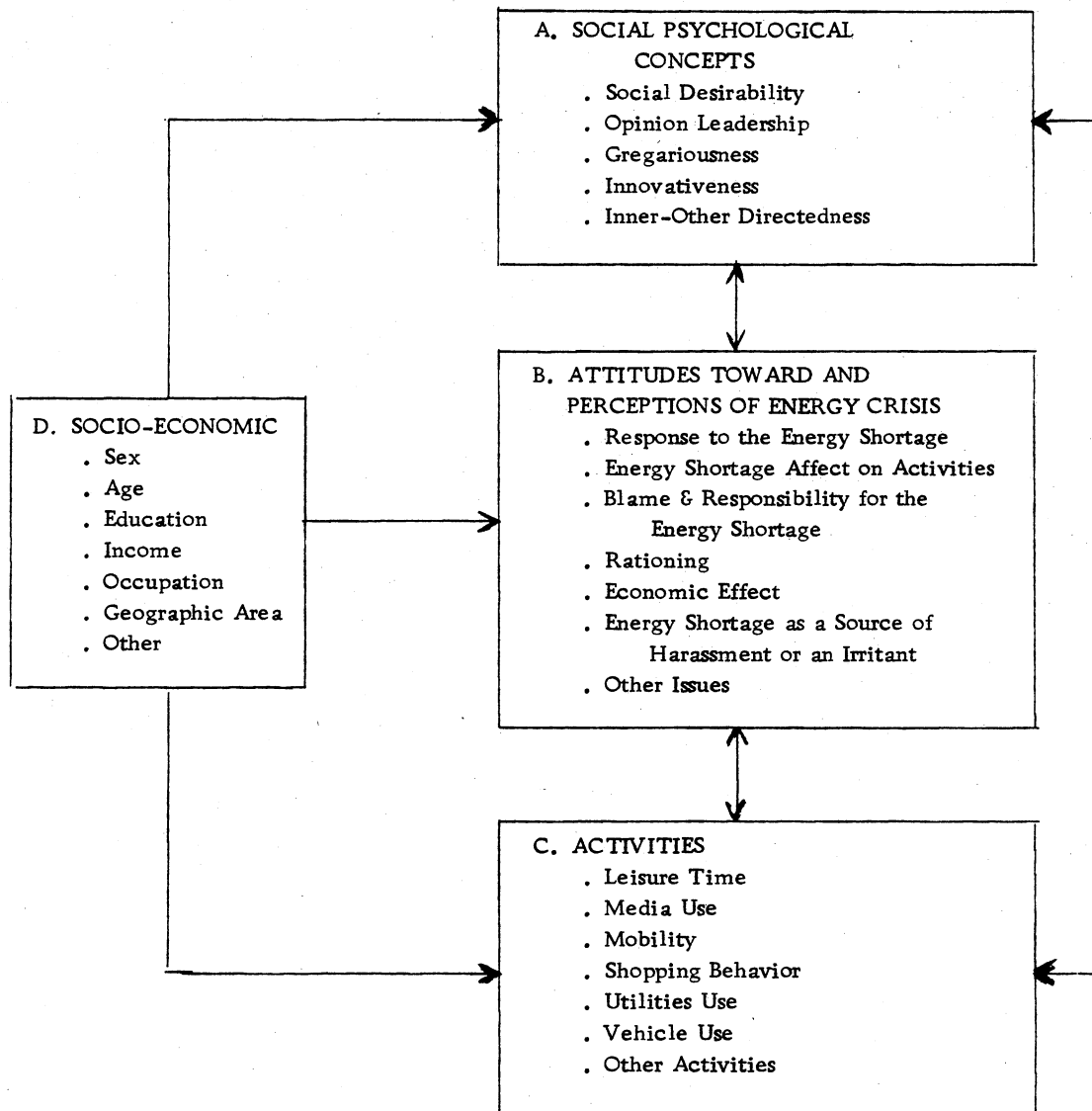


Figure 1. Overview of the energy crisis research project

said they were not affected by the energy shortage.

Specifically, the objectives of this paper are to examine among what consumers and to what extent:

- (1) changes in leisure time events were related to attitudes concerning the energy crisis
- (2) variations in shopping patterns were related to attitudes concerning the energy crisis

The relationship among the above objectives is portrayed in Figure 2 by the explanatory variables and the statements used to measure the two research areas.

#### Research Questions

In the absence of a theoretical foundation to lend structure to this

area of investigation, certain research questions were formulated. These questions, consistent with the prior stated objectives, involved analysis of responses to the four statements shown in Figure 2.

Agreement or disagreement with each of these four statements formed a set of dependent variables. Six areas of activities, interests, and opinions (AIO) statements served as independent variables along with selected socio-economic characteristics. The six areas of AIO statements, developed to be potentially related to the dependent variables included: attitudinal response to the energy shortage; expressed energy shortage affect on activities and other activities which may be affected by the shortage; responsibility for the energy shortage; rationing; economic effects; and the energy shortage as a source of harassment.

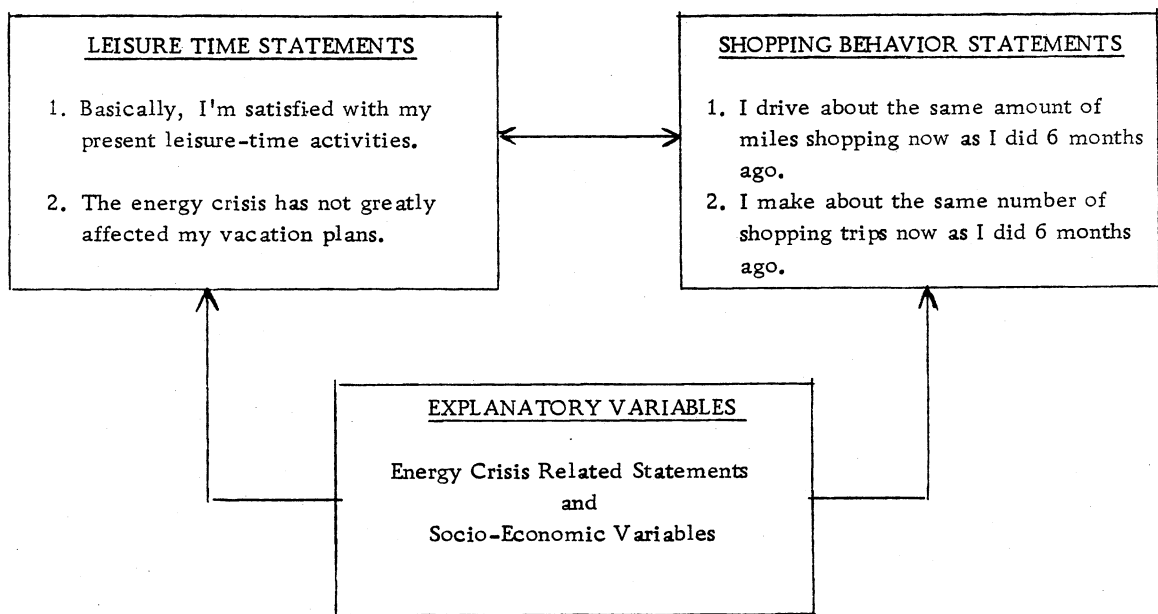


Figure 2. Interrelationships between the three research areas and the explanatory variables.

With this background the research questions became:

What percentage of the sample -

- is satisfied with present leisure-time activities?
- has not had vacation plans greatly affected by the energy crisis?
- drives about the same amount of miles shopping as before the energy crisis?
- makes about the same number of shopping trips as before the energy crisis?

Which variables best discriminate between those who -

- are satisfied and those who are not satisfied with present leisure-time activities
- have had and those who have not had vacation plans greatly affected by the energy crisis?

- . drive about the same and those who drive different amounts of miles shopping as before the energy crisis?
- . make about the same and those who make different numbers of shopping trips as before the energy crisis?

## Methodology

### Data Collection Instrument

Data to evaluate the research questions were gathered through a series of AIO statements scaled on a five-point scale from strongly agree to strongly disagree. As can be seen in Table 1, presented later, both the dependent variables and the independent variables with the exception of the socio-economic variables were scaled identically. Forty-five statements comprised the total range of AIO statements. For a more elaborate description of the questionnaire and the sample base discussed in the next section see Talarzyk and Omura (1974).

### Sample Base

The sample for this study involved 662 respondents, balanced to represent the United States population in terms of geographic dispersion, income, density of locale, degree of urbanization and age. The respondents completed a twelve-page questionnaire sent to 1000 members of the Consumer Mail Panel of Market Facts, Inc. The questionnaires were mailed to panel members during the first week of March, 1974.

### Data Analysis

To fulfill the objectives and respond to the research questions of this paper, three basic routines were performed on the data: frequency count, chi square analysis, and step-wise multiple discriminant analyses. The frequency count provided an aggregated overview of all the AIO items taken individually and focused on the first set of research questions. The chi square analysis allowed a basic investigation of the second set of research questions through an analysis of each of the four dependent variables versus each independent variable separately.

The discriminant analyses provided a more comprehensive evaluation of the second set of research questions by developing a profile of the characteristics that best aided in distinguishing between the two groups of each dependent variable. In addition, these analyses yielded a model for the prediction of group membership given knowledge of a respondent's answers to the differentiating statements.

For the chi square and discriminant analyses, the dependent variables were separately collapsed to form an agree group (strongly agree and agree respondents) and a disagree group (disagree and strongly disagree respondents). The undecided or no opinion group was eliminated from further analysis. When a given dependent variable was being studied the other three dependent variables were allowed to enter the analysis as assumed independent variables.

Hold-out samples were used to test the efficacy of the derived discriminant functions. The individuals for each hold-out sample were randomly drawn from each respective group prior to analysis. Approximately 25% of each original group went to the hold-out samples.

TABLE 1  
 Frequency Responses to All Variables  
 (In percentages -- Sample size = 662)

Variable Number	Variables	Strongly Agree	Agree Somewhat	Undecided or No Opinion	Disagree Somewhat	Strongly Disagree
<u>DEPENDENT*</u>						
1	Basically, I'm satisfied with my present leisure-time activities.....	29%	39%	8%	18%	7%
2	The energy crisis has not greatly affected my vacation plans.....	20	29	13	22	16
3	I drive about the same amount of miles shopping now as I did 6 months ago..	28	28	4	26	14
4	I make about the same number of shopping trips now as I did 6 months ago...	27	27	4	26	16
<u>INDEPENDENT</u>						
<u>Response to the Energy Shortage</u>						
5	I think that responding to the energy-saving pleas is patriotic.....	29%	38%	15%	12%	6%
6	I think that because of the energy crisis everyone should drive at 55 miles per hour or less.....	58	23	5	8	6
7	I think that responding to all the energy-saving pleas is fun.....	8	20	16	36	20
8	I've been ignoring the energy crisis.....	3	10	9	36	42
9	Everyone who lives near bus routes should take the bus if it gets them close to where they want to go.....	27	39	13	16	5
10	If at all possible, car pools should be formed for transportation.....	45	37	10	5	3
11	I'm not willing to sacrifice just so that all will have energy.....	4	8	13	36	39
<u>Energy Shortage Affect on Activities</u>						
12	The news about an energy crisis hasn't affected the way I live.....	10%	22%	7%	41%	20%
13	I try to include as many stores as possible in one shopping trip.....	61	25	4	6	4
14	I do most of my grocery shopping in one store.....	36	32	4	18	10
15	I wouldn't want to purchase groceries over the telephone.....	52	16	9	14	9
16	In recent years, I have been out of state at least 3 times a year.....	41	12	7	8	32
17	Television is my primary source of entertainment.....	30	33	4	20	14
18	I do a lot of traveling by car.....	30	33	4	20	14
19	I often drive around for pleasure or sightseeing.....	6	15	6	33	41
20	I often go out to dinner or the theater.....	8	22	5	29	36
21	I would like to take a long trip.....	44	26	12	8	10
22	I visit with my friends a lot.....	15	35	8	33	9
23	Everyone should take a vacation away from home at least once a year.....	54	29	8	7	2
24	I often work on a do-it-yourself project in my home.....	43	34	4	11	8
25	Shopping is a pleasurable activity, even if I don't buy anything.....	25	31	5	20	19
<u>Responsibility For Energy Shortage</u>						
26	The oil companies are making unreasonably high profits.....	59%	22%	13%	4%	2%
27	Oil companies should be forced to lower their prices.....	34	35	18	10	3
28	As soon as the Arab oil embargo is lifted, the energy shortages will disappear..	8	20	16	36	20
29	Gasoline stations are making unreasonably high profits.....	9	18	19	31	23
30	Even if the oil embargo were lifted, this country's energy problems will be a long way from being solved.....	32	45	13	7	3
31	The energy crisis is for real.....	20	30	19	20	4

Variable Number	Variables	Strongly Agree	Agree	Somewhat Undecided or No Opinion	Disagree	Somewhat Strongly Disagree
<u>Rationing</u>						
32	I expect nation-wide rationing to be put into effect before the year is out....	5%	10%	18%	32%	35%
33	Compared to the rest of the country, I think this area is getting its fair share of gasoline.....	24	38	16	14	8
<u>Economic Effect</u>						
34	The country is going to suffer from the effects of the energy crisis for several years to come .....	36%	40%	12%	9%	3%
35	The energy crisis is going to get worse before it gets better.....	15	32	26	22	5
36	I believe the country is headed into a serious recession.....	17	30	23	20	10
37	Prices will probably be a lot higher next year.....	53	34	7	4	2
38	I will probably have more money to spend next year than I have now.....	10	24	21	23	26
39	The energy crisis has greatly reduced my income.....	18	20	11	21	30
40	I am optimistic about the future of this country.....	38	36	15	8	3
41	I am optimistic about my personal future.....	30	36	16	15	3
42	If the energy shortage gets any worse, the country will be in bad shape.....	30	37	16	15	3
<u>Energy Shortage as a Source of Harassment or Irritant</u>						
43	Airlines would probably just as soon put wooden benches in their planes to get more people in.....	8%	8%	26%	25%	33%
44	A gasoline purchase limit is just a device to hassle me .....	11	18	21	27	23
45	I don't get the service I used to when I buy gasoline.....	44	25	7	14	10

\*DEPENDENT variables become INDEPENDENT variables when not being used as a dependent variable for analysis.

Although it will be apparent from Table 1 that the sizes of the two groups belonging to each dependent variable were not exactly equal, neither were the groups greatly disparate. The greatest disparity was less than a ratio of three to one. Still, in order not to be misled by the somewhat unequal group sizes, the proportional chance criterion was calculated for each dependent variable.

### Findings

#### Frequency Counts

Table 1 shows the frequency counts for all of the AIO statements used in this analysis. For the first set of research questions the following findings may be drawn: (1) some 68% of all respondents indicated that they were satisfied with their present leisure time activities, 25% said they were not satisfied; (2) the energy crisis apparently greatly affected the vacation plans of 49% of the respondents, 38% said they were not affected; (3) a total of 56% of the sample said they drive about the same amount of miles shopping now as they did before the energy crisis, 40% said they have changed the amount of driving for shopping; and (4) some 54% of the respondents stated that they make about the same number of shopping trips as they

TABLE 2  
Results from Chi Square and Discriminant Analyses

Variable Number	Predictor Variables	Dependent Variables (Groups)*							
		1 (X <sub>1</sub> ) (Y <sub>1</sub> )		2 (X <sub>2</sub> ) (Y <sub>2</sub> )		3 (X <sub>3</sub> ) (Y <sub>3</sub> )		4 (X <sub>4</sub> ) (Y <sub>4</sub> )	
1	Satisfied/leisure	X	X	0.80 <sup>b</sup>	1.14				
2	Not affect vacation		b	X	X	1.99 <sup>a</sup>	2.38		a
3	Miles shopping			0.66 <sup>a</sup>	1.26	X	X	0.34 <sup>a</sup>	2.24
4	Trips shopping				a	1.56 <sup>a</sup>	3.20	X	X
5	Patriotic		b	2.46	2.82	3.45 <sup>c</sup>	2.63		
6	Drive 55 MPH								
7	Fun		a						
8	Ignoring					4.51 <sup>a</sup>	3.37		a
9	Buses								
10	Car pools								
11	Not sacrifice					c			b
12	Not affect life	2.82 <sup>a</sup>	2.72	1.81 <sup>a</sup>	1.95	a		2.58 <sup>a</sup>	2.60
13	Stores/trip					a		1.93 <sup>a</sup>	1.22
14	Shop 1 store								c
15	Telephone groceries					c			
16	Been out of state								
17	TV entertainment								
18	Car travel								
19	Driving pleasurable					a			b
20	Dinner/theater								
21	Long trip					c			
22	Visit friends		b						
23	Vacation away	1.19	0.67						
24	Do-it-yourself	1.46 <sup>b</sup>	2.07	1.28	0.82	c			
25	Shopping pleasurable								c
26	Oil profits								
27	Lower oil prices				b				c
28	Embargo-disappear	2.37	4.22			c			
29	Station profits								
30	Embargo-solution								
31	Real								
32	Expect rationing								
33	Share gas		b		c	c			a
34	Suffer years								
35	Get worse								
36	Recession		c		c	2.26	2.62	1.95	1.67
37	Prices higher					1.67	1.19		
38	More money			2.86 <sup>a</sup>	3.09				
39	Reduced income								
40	Optimism-country								
41	Optimism-self	1.68 <sup>a</sup>	2.60						

Variable Number	Predictor Variables	Dependent Variables (Groups)*							
		1		2		3		4	
		(X <sub>1</sub> )	(Y <sub>1</sub> )	(X <sub>2</sub> )	(Y <sub>2</sub> )	(X <sub>3</sub> )	(Y <sub>3</sub> )	(X <sub>4</sub> )	(Y <sub>4</sub> )
42	If worse, bad								
43	Airlines								
44	Purchase limit	1.73	1.55	2.30 <sup>b</sup>	2.06				
45	Service-gas		<sup>b</sup>		<sup>c</sup>				<sup>c</sup>
46	Education-higher (+)			2.54	2.75	2.36	2.47	2.49	1.69
47	Household size-larger (+)						<sup>c</sup>		<sup>b</sup>
48	Income-higher (+)			2.08	1.90		<sup>a</sup>		
49	Age-older (+)	2.73 <sup>a</sup>	2.50				<sup>b</sup>	2.80	2.28

\*Chi square analysis performed on the above row by column variables via 5 x 2 tables except for row variables 47 and 48 which were done via 9 x 2 tables. The numbers shown represent standardized discriminant coefficients which were significant at the 0.05 level.

<sup>a</sup>Chi square test significant at 0.001 level

<sup>b</sup>Chi square test significant at 0.01 level

<sup>c</sup>Chi square test significant at 0.05 level

made before the energy crisis, 42% said they have changed the number of shopping trips.

#### Chi Square Pairwise Variable Tests

As a part of the analysis of the second area of research questions, chi square runs were made. The following paragraphs in connection with the information in Table 2 briefly describe the research findings in this area.<sup>4</sup> The findings are discussed in terms of the general areas of investigation which were included in the study rather than each individual independent variable in order to conserve space.

Variable 1 (Satisfied with leisure activities): Eleven variables were determined to be significantly related to satisfaction with leisure activities. As seen in Table 2, each of the areas except responsibility for the energy shortage is found to be related to Variable 1 (read across the columns as Independent Variable 1, with the agree group as X<sub>1</sub> and disagree group as Y<sub>1</sub>) through at least one of the area's subcomponents.

Variable 2 (Energy crisis not affecting vacation plans): Ten variables were found related to Variable 2 but not all of the dependent areas were represented. None of the attitudinal response variables were determined to be related to Variable 2, nor were any socioeconomic variables related.

Variable 3 (Miles shopping unchanged in past 6 months): Although seventeen variables were discovered to be related to Variable 3, the entire area of anticipation of economic conditions was unrelated.

Variable 4 (Number of shopping trips unchanged in past 6 months): Thirteen variables were found to be related to Variable 4 with the area of anticipation of economic conditions not related.

While the chi square analyses indicated that relationships existed between the two groups of variables in the sample under study, only a gross association could be stated. Information useful to a consumer behavioralist concerned with the energy shortage was obtained more concisely through discriminant analyses.

#### Profiles through the Multiple Discriminant Analyses

Before profiling the groups under each of the five dependent variables, it should be noted that while significant discriminating variables were uncovered, few of the coefficients were to any great extent different relative to the corresponding variable in the opposing group. Thus, while the written descriptions below tend to draw out relative differences, the actual coefficients of each of the two groups should be observed also in Table 2. The number in parentheses following each of the characteristics below refer to the corresponding variable number from which the descriptions were drawn.

#### Variable 1 (Satisfied with leisure activities)

Those more likely to be satisfied:

- reported their lifestyles to be more affected by the energy crisis (12)
- reported less often that everyone should vacation away from home (23)
- worked more on home projects (24)
- held greater belief that the energy crisis would disappear with the lifting of the embargo (28)
- were more optimistic about their personal future (41)
- held less belief that the purchase limit was a hassling device (44)
- older (49)

The attitudinal response to the energy crisis and the rationing areas were not included as sources of discriminating characteristics under Variable 1.

#### Variable 2 (Energy crisis not affecting vacation plans)

Those less likely to be affected:

- more satisfied with their leisure-time activities ( 1)
- drive same number of miles shopping presently as before the crisis ( 3)
- held greater belief that responding to energy-saving pleas was patriotic ( 5)
- reported their lifestyles to be less affected by the energy crisis (12)
- worked less on home projects (24)
- held greater belief that they would have more money to spend next year (38)



- held less belief that the purchase limit was a hassling device (44)
- less educated (46)
- had higher income (48)

The areas of responsibility for the energy shortage and rationing were not represented in terms of contributing significant discriminating characteristics to Variable 2.

#### Variable 3 (Miles shopping unchanged in 6 months)

Those who have not altered the number of miles shopping:

- less likely to have their vacation plans affected by the crisis ( 2)
- reported taking the same number of shopping trips presently as before the crisis ( 4)
- held less belief that responding to energy-saving pleas was patriotic ( 5)
- reported themselves to be ignoring the energy crisis less ( 8)
- held more belief that the country was recession-bound (36)
- held less belief that prices would be higher next year (37)
- less educated (46)

The only areas which led to discriminating between those who have not altered the number of miles shopping and those who have, include attitudinal response to the energy shortage, economic effect, and socio-economic characteristics.

#### Variable 4 (Number of shopping trips unchanged in 6 months)

Those who have not altered the number of shopping trips:

- drive same number of miles shopping presently as before the crisis ( 3)
- reported their lifestyles to be less affected by the energy crisis (12)
- tried less to include many stores in one shopping trip (13)
- held less belief that the country was recession-bound (36)
- more educated (46)
- older (49)

Three areas contributed to discriminating between the two groups of this dependent variable: energy shortage affect on activities, economic effect, and socio-economic characteristics.

#### Validation of the Discriminant Functions

In order to determine how well the variables used in the discriminant functions discriminate, the proportional chance criterion  $C_{pro}$  was calculated. This criterion serves as an index of how well both groups are correctly identified (Morrison, 1969). The criterion  $C_{pro}$  indicates the probability an individual may be correctly classified by chance. If the percentage of the total number correctly classified exceeds  $C_{pro}$ , then the model derived is better than what might have occurred through chance. Table 3 provides the critical figures.

TABLE 3  
Validation of Discriminant Functions

Variable Number	Variable	Per Cent Correct			C <sub>pro.</sub> (%)
		Total (N)	X <sub>i</sub> (n)	Y <sub>i</sub> (N-n)	
1	Satisfied/leisure				
	Original groups	67.3% (459)	67.6% (336)	66.7% (123)	60.8%
	Test groups	65.1 (152)	65.8 (111)	63.4 (41)	60.6
2	Not affect vacation				
	Original groups	68.3 (432)	68.3 (246)	68.3 (186)	51.0
	Test groups	57.3 (143)	59.3 (81)	54.8 (62)	50.9
3	Miles shopping				
	Original groups	81.1 (477)	81.9 (276)	80.1 (201)	51.2
	Test groups	83.4 (157)	81.3 (91)	86.4 (66)	51.3
4	Trips shopping				
	Original groups	82.8 (476)	86.0 (264)	78.8 (212)	50.6
	Test groups	79.7 (158)	86.4 (88)	71.4 (70)	50.6

(1)  $C_{pro.} = a^2 + (1-a)^2$  where  $a$  is the proportion of the sample belonging to one of the two subgroups.

(2)  $X_i$  - the agreement groups;  $Y_i$  - the disagreement groups.

For both the original sample and the holdout sample, the derived discriminant functions showed high levels of correct classifications. The 65.1% correct prediction of those satisfied/unsatisfied with their leisure time activities and the 57.3% correct prediction of those reporting the energy shortage not affecting/affecting their vacation plans both exceeded their respective proportional chance criterion reasonably well.<sup>5</sup> The last two groups, those reporting no change/change in miles shopping and shopping trips since the energy shortage were predicted correctly exceptionally well, 83.4% and 79.7%, respectively, far exceeding their respective  $C_{pro.}$ .

#### Discussion

The findings demonstrated that relationships indeed existed between consumer attitudes toward the energy shortage and certain leisure time activities and shopping patterns. The relationships were evaluated through the frequency analysis, chi square analysis, and discriminant analysis. In the frequency count, there was by no means unanimity in the direction of responses to the dependent variable statements, suggesting that consumers were differentially affected by or differentially responded to the energy shortage. The chi square analysis indicated certain general associations between the selected variables, while the discriminant analysis pinpointed characteristics of those reporting basic satisfaction with leisure time pursuits and the reported affect of the energy shortage on leisure activities and shopping patterns. The discriminant analysis indicated that the dichotomous groups could be differentiated, and even better, that membership in the groups could be successfully predicted on the basis of the discriminating variables found to be significant in the respective models.

In examining the profiles reported above, a number of interesting

observations may be made. Those more likely to be satisfied with their leisure time activities reported their lifestyles to be more affected by the energy shortage, but apparently the satisfaction arose or remained because their leisure activities did not include a great deal of energy consumption. Also, a general degree of optimism was present.

The respondents who indicated that the energy shortage was not going to affect their vacation plans apparently did so because of financial well-being. This conclusion is suggested by the importance of these financial variables (38 and 48) in the discriminant model seen in Table 2. At the same time it is curious that these respondents held greater belief that responding to energy-saving pleas was patriotic.

Perhaps the explanation is that the life styles of this group do not include extensive traveling on vacations, or, rather, that they simply ignored the energy-saving pleas despite recognizance of the obvious patriotic response. If the latter explanation is more correct, and the complete profile tends to indicate this, patriotic or patriotic-types of pleas through the media apparently will not induce a large segment of people, 38% of this sample, to reduce energy consumption. Should the country face another energy crisis or suffer further deterioration in energy supply, further research appears necessary to determine the type of appeal that would be positively received and acted upon other than the past choice of patriotic-oriented pleas.

The group that indicated no change in the numbers of shopping miles driven and the group with no change in the number of shopping trips made exhibited the expected similar responses to the classification variables. As groups, in general, they reported less influence on their life styles by the energy shortages. There were some variations in certain independent variables (belief that the country was recession-bound) and socio-economic variables (age and education). But the differences were slight.

While this investigation has provided only an abbreviated examination of the relationships between consumer attitudes toward the energy shortage and certain activities, it will hopefully serve as part of a composite research effort. Such a larger effort is required for consumer behavioralists to fully understand the spectrum of interrelationships involving this contemporary social issue.

As part of on-going research, additional analyses of the larger data set from which this paper is based will aid in a further understanding of the impact of major social issues on consumers. The issues of the energy shortage and the presidential impeachment will be examined through socio-psychological theories of opinion leadership, innovativeness, social character, gregariousness, and social desirability.

#### FOOTNOTES

1. The authors gratefully acknowledge research support provided by the Fred B. and Mabel Dean Hill Fund and Consumer Mail Panels of Market Facts, Inc.
2. Glenn S. Omura is a Research Associate and Doctoral Candidate in the College of Administrative Science at The Ohio State University.

3. W. Wayne Talarzyk is Associate Professor of Marketing of the College of Administrative Science, The Ohio State University.
4. Because of space limitations the individual contingency tables are not shown here. Detailed results are available upon request from the authors.
5. For a comparison of how well discriminant models in another context predicted, see Ostlund (1974).

#### REFERENCES

- Morrison, D. G. On the interpretation of discriminant analysis. Journal of Marketing Research, 1969, 6, 156-163.
- Ostlund, L. E. Perceived innovation attributes as predictors of innovativeness. Journal of Consumer Research, 1974, 1, 23-29.
- Talarzyk, W. W. & Omura, G. S. Consumer attitudes toward and perceptions of the energy crisis. Proceedings, Fall Conference, American Marketing Association, 1974.



CONSUMER SATISFACTIONS FROM LEISURE TIME PURSUITS<sup>1</sup>

Douglass K. Hawes<sup>2</sup>  
University of Wyoming

W. Wayne Talarzyk<sup>3</sup>  
The Ohio State University

Roger D. Blackwell<sup>4</sup>  
The Ohio State University

Leisure time pursuits are increasingly becoming a topic of major concern to scholars and practitioners in many different disciplines. This study was an attempt to define the patterns of usage of leisure time pursuits of Americans, and to analyze the satisfactions derived from leisure time activities. Findings are presented in a series of tables.

The study found that it is possible to determine the satisfactions which consumers derive from their leisure time pursuits. It is also possible to group leisure time pursuits and to group satisfactions into interpretable clusters. Some discussion is presented on the utility of these findings to businessmen, and to planners in the public sector.

Leisure time pursuits are increasingly a topic of major concern to scholars in a variety of disciplines. Marketing practitioners are attracted by the massive amounts of money, variously estimated between \$80-\$150 billion devoted to leisure time goods and services. Politicians and resource planners are concerned with the rapidly increasing demands on public recreation facilities and extensive shortages likely to occur in certain leisure facilities. Sociologists and social psychologists may be increasingly interested in the need to understand mental health in terms of time usage patterns and in quality of life indicators. Consumer behavior theorists realize that a comprehensive explanatory model of consumer choice must include analysis of time budget influences as well as money budget influences.

#### Overview of the Study

The purpose of this study was to describe leisure time pursuits of Americans and to analyze the satisfactions derived from leisure time activities. The study encompasses additional research questions, but in this paper, the following questions are investigated:

1. What are the major leisure time pursuits of American consumers?
2. Are there identifiable "satisfactions," or perceived felt benefits, which people (participants) derive from leisure time pursuits?
3. If there are identifiable "satisfactions," do these "satisfactions" cluster or group together in some manner based upon particular leisure time pursuits?

4. Do people typically engage in definable, fairly distinct, clusters of leisure time pursuits to the exclusion of other pursuits? If so, how strongly are the pursuits within a cluster related? Can these clusters be considered reachable, viable market segments?

### Research Hypotheses

In addition to the descriptive portions of this study, the research questions can be restated as research hypotheses:

- H<sub>1</sub>: There are definable "satisfactions" which people derive from leisure time pursuits.
- H<sub>2</sub>: Leisure time pursuits and satisfactions can be clustered into distinct groups.

### Definitions of Terms

The following definitions were used in this study:

1. Leisure time--Time perceived by the respondent as not obligated a priori to work, work-related activities, life maintenance activities, routine family duties and responsibilities, and routine social and civic responsibilities.
2. Leisure time pursuits--Those endeavors, either passive or active, which people undertake during their leisure time. "Pursuit" is used rather than activity because it does not have the connotative restriction of "activeness."
3. Groups of participants--Aggregations of individuals defined by their leisure time life style and pursuits. A verbal profile of groups of individuals built up by relating an individual and his or her situational parameters either to favorite leisure time pursuits or those pursuits participated in the year prior to the study.
4. Satisfactions--The meanings or significance which leisure-time pursuits hold for the respondent, as perceived by the respondent. These meanings can be viewed as perceived psychological "outputs" in a model of decision making or benefits from participating in a pursuit.

These meanings or benefits are embodied in a list of statements (hereafter called "satisfactions statements") of unitary affective content. The respondent indicates the relevance of each satisfaction to him or her (as an output from participation in a favorite pursuit) by rating each statement on a five-point "importance" scale.

### Research Methodology

The research methodology used in this study involved collecting data on a fairly comprehensive group of leisure time oriented variables from a representative cross-section of the American population. The study was designed to collect sufficient data to provide analysis of a wide range of leisure time related topics.

### Data Collection

The instrument used to collect data was a questionnaire containing, in part, the following information:

1. respondent's favorite leisure-time pursuits (three pursuits selected from a list of 50).
2. "satisfactions" statements rated on a 1-5 scale from "very important" to "not important."
3. respondent's participation in leisure-time pursuits in the past, and during 1972.
4. respondent's rating (from "strongly agree" to "strongly disagree") of 87 AIO (activity, interest, opinion) statements.
5. demographic variables including:
  - a. respondent's age
  - b. respondent's religion
  - c. respondent's education
  - d. total family income
  - e. household size
  - f. occupation of husband
  - g. rural/urban location of household.

Other information such as time budgets, perceived time/money tradeoffs, and media preferences, was also collected and will be presented in additional reports when analysis is completed.

A pre-test was conducted among students to determine the clarity of the questions and a secondary pre-test of the questionnaire was conducted among 85 households in the Columbus area. Based on this pre-testing it was decided to use a five-point scale instead of a seven-point scale on AIO questions in order to minimize respondent fatigue. Seventeen statements were also eliminated from the final survey instrument because they failed to meet various response criteria such as discriminatory ability. Numerous other modifications were made based upon the pre-test results. The final pre-test was conducted among 100 households throughout the United States using the same sampling procedure described below.

### Sampling Procedure and Response Rate

Two questionnaires were mailed to 1000 households in May, 1973 by Market Facts, Inc., from their panels of 45,000 households. Questionnaires for males and for females were sent to households whose names had recently been added to the panel, and were sent to households with demographic characteristics proportional to demographic characteristics for the United States. A total of 603 usable female and 512 usable male questionnaires were returned. There were 23 additional female and 96 additional male questionnaires which were returned incompletely answered. While these could have been used for specific questions, the following analyses were performed only on the completely returned questionnaires. The 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 similar to the total United States.

### Research Findings

The findings concerning leisure time pursuits and the satisfactions associated with them are presented in Tables 1 through 10. A brief commentary is presented below on the findings of each of these tables.

#### Leisure Time Pursuits of American Consumers

The most popular, and hence most frequently pursued, leisure time pursuits of American consumers are described in Table 1. Among women consumers, the most frequent leisure time pursuit is "listening to music from records, tapes, FM or AM radio" and was engaged in by almost 90 percent of the sample women during 1972. The second and third most popular pursuits were "visiting with friends, partying" and "reading a book for pleasure," and were engaged in by 89.2 percent and 88.7 percent of the sample respectively.

It is interesting to note that the five most popular pursuits among the women are basically indoor pursuits. It is not until one looks at the sixth and seventh ranked pursuits ("attending movies" (84.2 percent) and "driving around for pleasure or sightseeing" (84.1 percent)) that truly outside-the-home pursuits appear. Finally, ten different pursuits were engaged in by more than 80 percent of the sample women, and of these ten, seven are basically indoor, around-the-home pursuits.

A somewhat different pattern is evident among men consumers. For male members of the sample, the most frequent leisure time pursuit during 1972 was "visiting with friends, partying"--engaged in by 85.4 percent of the sample. The second and third most popular pursuits were "driving around for pleasure or sightseeing" (85.2 percent), and "listening to music from records, tapes, FM or AM radio" (83.2 percent).

Among the top five pursuits for the males, only three are basically indoor pursuits. Six of the top ten pursuits are in the indoor, around-the-home category. Finally, only the four most popular pursuits were engaged in by more than 80 percent of the sample, while the tenth most popular attracted only 63 percent of the men. This implies a more diffuse, extensive set of leisure time interests among men than among women. Women appear to be somewhat more homogeneous and intensive in their leisure time interests.

Table 2 (females) and 3 (males) present the basic demographic characteristics of participants in the 42 most popular leisure time pursuits. The figures given are rounded percentages and may be compared with the percentage distributions in the total sample which are listed in Table 4 below. The volume of data in these two tables does not readily lend itself to a verbal summary analysis, so discussion will be limited to several selected relationships.

TABLE 1

Percent of Sample Females and Males Participating  
In Each of the 47 Defined Leisure-Time Pursuits  
At Least Once During 1972<sup>a</sup>

Leisure-Time Pursuits	Females		Males	
	(N <sub>tot</sub> = 603)	(Rank)	(N <sub>tot</sub> = 512)	(Rank)
Attending concerts or plays	50.1%	(20)	37.1%	(25)
Camping by trailer, camper, or motor home	23.7	(35)	24.0	(32)
Camping by tent	21.1	(37)	25.6	(30)
Canoeing, rowing, rafting	14.3	(40)	22.1	(34)
Fishing or hunting	40.1	(24)	58.8	(13)
Golf	10.4	(41)	23.4	(33)
Hiking, backpacking, nature study	19.2	(38)	21.9	(35)
Horseback riding	22.1	(36)	19.9	(37)
Ice skating, roller skating	29.7	(32)	24.2	(31)
Picnicking	81.8	(9)	75.0	(6)
Power boating, water skiing, scuba diving	25.2	(34)	30.5	(26)
Sailing	5.8	(44)	7.6	(42)
Snowmobiling	8.0	(42)	8.4	(41)
Snow skiing	5.6	(45)	5.9	(44)
Swimming	60.5	(16)	6.3	(43)
Tennis	15.9	(39)	16.6	(39)
Square-dancing or other organized dances	30.5	(31)	20.5	(36)
Attending movies	84.2	(6)	82.4	(4)
Attending sporting events (as a spectator)	57.4	(17)	64.8	(9)
Automobile modifications, tune-ups, etc.	7.1	(43)	44.3	(19)
Playing basketball, football, baseball, softball, volleyball, handball	31.0	(30)	45.3	(18)
Bicycling	43.3	(22)	38.9	(24)
Bingo, bridge, or similar card games; "boxed" or board games such as Monopoly, Scrabble, etc.	80.6	(10)	67.2	(8)
Bowling	43.3	(22)	43.9	(20)
Chess, checkers, backgammon	36.5	(27)	39.8	(23)
Church-related activities	55.5	(18)	43.4	(21)
Collecting coins, stamps, bottles, shells, nic-nacs, antiques, etc.	40.6	(23)	29.3	(27)
Fixing up the house, remodeling, making repairs	63.7	(15)	73.8	(7)
Gardening, lawn care, landscaping	68.3	(13)	75.0	(6)
Pool or billiards; table-tennis	37.8	(26)	56.3	(15)
Volunteer, community, school, youth group, or charitable organization work	43.4	(21)	27.3	(29)
Racing or rallying (car or bike)	3.5	(46)	8.4	(41)
Creative crafts or handicrafts such as painting, drawing, sculpturing, sewing, knitting, embroidery, candle making, etc.	83.9	(8)	18.4	(38)
Driving around for pleasure or sightseeing	84.1	(7)	85.2	(2)
Exercising, jogging, visiting a health spa	38.5	(25)	28.3	(28)
Job-related reading or study	34.5	(28)	48.6	(17)
Listening to music from records, tapes, FM or AM radio	89.9	(1)	83.2	(3)
Photography, taking pictures	70.5	(12)	61.7	(12)
Playing the piano, organ or other musical instrument for pleasure	28.9	(33)	15.2	(40)
Playing with children	84.4	(5)	77.1	(5)
Reading a book for pleasure	88.7	(3)	62.9	(10)
Reading the Bible	66.7	(14)	42.2	(22)
Visiting a bar or club	50.6	(19)	61.9	(11)
Visiting with friends, partying	89.2	(2)	85.4	(1)
Walking for pleasure	71.8	(11)	58.4	(14)
Woodworking, metalworking, furniture refinishing, home workshop projects	31.5	(29)	53.7	(16)
Writing letters, doing crossword puzzles	85.9	(4)	39.8	(23)

Notes:

<sup>a</sup>There were 47 defined leisure-time pursuits grouped into 3 categories - each with an "Other" alternative.

TABLE 2  
Basic Demographic Characteristics of Female  
Participants in the "Most Popular" Leisure-Time  
Pursuits During 1972 - By Percent

Education	Concerts or plays (302) <sup>a</sup>	Camping by trail-er, etc. (143)	Camping by tent (127)	Canoeing, rowing, etc. (86)	Fishing or hunting (242)	Golf (63)	Hiking, backpacking, etc. (116)	Horseback riding (133)	Ice or roller skating (179)	Picnicking (493)	Power boating, water skiing (152)	Swimming (365)	Tennis (96)	Square-dancing, etc. (134)	Attending movies (508)	Attending sporting events (346)	Playing basketball, football, etc. (18)	Bicycling (261)	Bingo, bridge, etc. (486)	Bowling (261)	Chess, checkers, etc. (220)	Church-related activities (335)	Collecting coins, stamps, etc. (245)	Fixing up the house (384)	Gardening, lawn care, etc. (412)	Pool or table tennis (228)	Volunteer, community, school (262)	Creative crafts or handicrafts (506)	Driving around for pleasure (507)	Exercising, jogging, etc. (232)	Job-related reading (208)	Listening to music (542)	Photography (425)	Playing the piano, organ, etc. (174)	Playing with children (509)	Reading a book for pleasure (535)	Reading the Bible (402)	Visiting a bar or club (305)	Visiting with friends (538)	Making for partying (433)	Woodworking, metal-working, etc. (190)	Writing letters, crossword puzzles (518)	
Some or no High School (122) <sup>b</sup>	28	26	21	9	45	4	13	22	28	77	21	51	7	19	72	47	30	34	72	38	30	53	37	61	64	30	26	80	80	29	18	84	69	17	82	82	66	40	80	80	65	20	81
High School Graduate (268)	44	24	23	17	43	10	21	21	30	83	24	62	12	19	86	57	29	43	81	44	37	55	40	62	67	42	43	86	84	36	25	92	71	27	85	89	65	50	91	72	31	87	31
Some College (145)	66	23	18	10	36	17	16	25	32	80	30	63	22	29	88	66	34	51	87	46	40	58	46	67	74	34	52	86	84	46	52	91	70	41	85	92	72	54	90	74	40	98	48
College Graduate or Post-graduate (65)	85	18	20	25	26	11	34	20	29	91	28	69	34	25	91	62	35	45	80	43	39	55	38	68	71	46	58	77	94	51	62	92	74	31	88	95	65	95	80	37	86	37	
Total Family Income	39	21	24	13	48	5	16	18	29	78	19	50	14	15	71	50	24	33	66	29	33	53	30	38	55	29	31	83	75	24	18	80	61	21	79	84	63	31	78	65	24	76	
\$ 4,000 - \$ 5,999 (80)	40	19	27	11	42	7	24	34	33	91	24	69	18	27	82	55	46	48	81	47	40	63	43	74	73	52	35	82	90	42	39	97	75	30	86	88	71	55	89	83	36	89	
\$ 6,000 - \$ 7,999 (88)	38	31	27	13	39	7	20	29	32	80	29	58	13	24	82	57	35	50	76	48	33	51	37	71	69	33	32	83	82	42	26	85	73	26	86	83	67	46	89	69	25	85	
\$ 8,000 - \$ 9,999 (84)	40	20	21	14	46	9	12	17	29	81	21	62	11	17	86	57	28	46	88	45	43	57	44	61	67	34	50	85	82	36	32	87	65	27	87	87	70	39	86	62	29	86	
\$10,000 - \$11,999 (94)	53	27	15	15	34	12	24	24	29	78	26	63	21	21	87	51	29	45	78	50	38	51	41	67	64	38	47	89	84	37	32	91	75	31	90	93	67	54	91	75	35	88	
\$12,000 - \$14,999 (105)	72	24	16	17	36	17	19	16	28	83	30	60	17	26	91	67	28	40	88	41	34	58	45	67	76	39	55	82	88	45	49	95	72	34	80	93	65	65	97	74	36	89	
\$15,000 and up (152)	49	28	35	20	57	12	29	43	55	92	37	84	28	25	98	72	54	65	92	69	53	51	42	64	61	58	32	90	95	61	36	96	79	32	95	94	61	69	96	86	39	91	
Age	48	21	30	16	41	9	24	34	38	84	31	76	25	21	95	56	51	59	87	54	48	51	36	71	76	54	41	88	90	46	37	96	79	30	91	91	63	61	91	77	36	84	
Less than 25 (127)	59	28	20	15	37	8	19	16	32	85	27	67	14	29	89	66	28	57	83	50	37	63	45	72	76	38	54	85	83	36	41	97	71	32	88	91	66	54	91	72	30	84	
25 - 34 (140)	53	21	9	13	34	13	14	12	17	78	20	47	10	25	80	56	12	32	78	30	29	62	53	62	70	29	55	80	81	32	40	90	71	29	82	89	76	49	88	66	28	90	
35 - 44 (92)	46	21	10	8	31	11	10	4	6	71	11	29	2	14	61	40	6	9	64	14	16	55	32	53	60	9	41	77	71	16	21	74	54	24	68	80	69	23	81	58	24	81	
45 - 54 (104)	55	21	10	8	31	11	10	4	6	71	11	29	2	14	61	40	6	9	64	14	16	55	32	53	60	9	41	77	71	16	21	74	54	24	68	80	69	23	81	58	24	81	
55 and up (140)																																											

<sup>a</sup>Number in parentheses is number of participants.  
<sup>b</sup>Number in parentheses is number of respondents in that demographic category.  
<sup>c</sup>Cell value is percentage of respondents in that demographic category who engaged in the pursuit at least once during 1972.

TABLE 3  
Basic Demographic Characteristics of Male  
Participants in the "Most Popular" Leisure-Time  
Pursuits During 1972 - By Percent

Education	Concerts or plays (190) <sup>a</sup>	Camping by trail-er, etc. (123)	Camping by tent (131)	Canoeing, rowing, etc. (113)	Fishing or hunting (301)	Golf (120)	Hiking, back- packing, etc. (112)	Horseback riding (102)	Ice or roller skating (124)	Picnicking (384)	Power boating, water skiing (156)	Swimming (320)	Tennis (85)	Square-dancing, etc. (105)	Attending movies (422)	Attending sporting events (332)	Auto modifications, etc. (227)	Playing basketball, football, etc. (232)	Bicycling (199)	Bingo, bridge, etc. (344)	Bowling (225)	Cheers, checkers, etc. (204)	Church-related activities (222)	Collecting coins, stamps, etc. (150)	Fixing up the house (378)	Gardening, lawn care, etc. (384)	Pool or table tennis (288)	Volunteer, community, school (140)	Creative crafts or handicrafts (94)	Driving around for pleasure (436)	Exercising, jogging, etc. (145)	Job-related reading (249)	Listening to music (429)	Photography (316)	Playing the piano, organ, etc. (78)	Playing with children (395)	Reading a book for pleasure (322)	Reading the Bible (216)	Visiting a bar or club (317)	Visiting with friends (437)	Walking for pleasure (299)	Woodworking, metal-working, etc. (275)	Writing letters, word puzzles (204)		
Some or no High School (137)	18%	27	28	17	64	10	17	19	20	72	27	52	4	22	71	53	41	29	30	59	31	35	38	28	68	71	47	18	15	80	19	16	34	75	46	9	75	51	39	76	80	47	45	23	
High School Graduate (164)	33	23	30	24	63	24	24	25	27	79	32	67	13	22	88	68	51	48	37	68	49	38	45	28	21	88	28	59	26	21	88	28	45	87	65	20	74	59	43	65	85	56	61	38	
Some College (109)	44	24	21	24	57	27	19	23	26	71	29	69	22	14	81	66	44	51	45	70	50	44	51	35	76	74	57	28	19	83	38	55	87	70	19	79	73	47	58	87	62	52	48		
College Graduate or Post-Graduate (95)	66	22	19	25	47	39	31	10	25	78	34	66	35	24	92	76	38	58	47	75	45	47	42	27	80	78	62	44	17	91	37	72	88	71	14	85	79	43	72	95	79	58	58		
Total Family Income																																													
\$ 4,000 - \$ 5,999 (67)	19	27	27	15	61	10	16	25	18	69	21	43	9	18	64	48	42	40	31	63	31	30	49	24	51	66	43	19	19	73	18	28	75	46	13	69	49	45	42	78	52	48	31		
\$ 6,000 - \$ 7,999 (80)	29	19	35	25	63	19	28	23	26	80	35	64	21	18	78	63	45	46	32	61	48	41	48	26	70	75	60	20	24	89	25	43	90	58	25	83	60	53	60	85	59	64	40		
\$ 8,000 - \$ 9,999 (72)	22	36	28	26	58	25	21	26	24	82	38	71	13	25	75	68	46	53	47	64	46	36	43	32	78	74	58	33	19	86	31	46	86	63	13	78	64	36	65	83	61	57	44		
\$10,000 - \$11,999 (82)	31	23	22	20	62	23	18	13	21	74	22	63	15	12	85	62	48	43	40	66	43	43	37	28	73	74	50	16	13	83	29	44	79	59	11	74	55	38	59	78	52	43	28		
\$12,000 - \$14,999 (90)	46	27	23	24	62	29	26	20	31	74	31	67	20	19	89	66	49	47	48	76	52	53	39	36	77	77	63	36	16	90	30	53	84	69	11	86	71	47	68	88	63	49	48		
\$15,000 and up (121)	60	17	22	22	50	29	22	16	24	72	34	64	19	28	93	75	39	44	35	70	42	35	46	29	85	80	59	35	19	87	33	65	86	69	17	74	71	37	70	94	60	56	44		
AGE																																													
Less than 25 (64)	44	25	44	31	78	20	37	52	53	81	47	83	28	31	92	69	67	73	59	80	63	53	45	28	70	66	83	16	20	88	44	42	94	70	22	88	64	42	78	92	66	66	39		
25 - 34 (150)	34	29	35	32	64	27	27	27	33	83	44	79	29	22	93	66	61	73	49	79	64	48	39	34	77	77	74	24	24	89	36	61	88	69	19	87	61	41	73	85	66	58	42		
35 - 44 (63)	46	24	21	19	56	30	16	13	19	76	24	68	16	29	92	70	40	49	44	65	49	46	44	27	78	78	57	40	16	86	30	57	81	59	13	84	62	35	75	84	51	56	32		
45 - 54 (93)	39	27	22	23	55	23	13	14	70	26	61	9	17	77	65	37	30	37	61	37	38	43	32	79	83	52	30	10	87	26	52	87	59	11	74	68	45	58	87	56	51	42			
55 and up (142)	32	17	13	9	49	18	12	6	11	66	15	35	4	13	66	60	23	12	18	54	17	24	47	24	68	70	28	18	79	14	33	74	53	12	61	62	45	39	82	52	45	40			

<sup>a</sup>Number in parentheses is number of participants

<sup>b</sup>Number in parentheses is number of respondents in that demographic category.

<sup>c</sup>Cell value is percentage of respondents in that demographic category who engaged in the pursuit at least once during 1972.

TABLE 4

Percentage Distribution of Sample Respondents on the  
Demographic Variables of Education, Income and Age

	Females (N <sub>tot</sub> =603)	Males (N <sub>tot</sub> =512)
<u>Education</u>		
Some or no high school	20%	27%
High school graduate	44	32
Some college	24	21
College graduate or post-graduate	11	19
<u>Total Family Income</u>		
\$4,000-\$5,999	13%	13%
\$6,000-\$7,999	15	16
\$8,000-\$9,999	14	14
\$10,000-\$11,999	16	16
\$12,000-\$14,999	17	18
\$15,000 and up	25	24
<u>Age</u>		
Less than 25	21%	13%
25-34	23	29
35-44	15	12
45-54	17	18
55 and up	23	28

Among the females (Table 2), it is evident that golf and tennis are played by the better educated and wealthier women. Disproportionately more women less than 34 years of age prefer tennis, while golf, on the other hand, tends to be preferred by those less than 25 and over 45. By contrast, picnicking, attending movies, playing bingo, bridge or other card games, driving around for pleasure, listening to music, playing with children, reading a book, visiting with friends, and writing letters or doing cross-word puzzles have a broad appeal to all categories of education, income, and age.

Several interesting patterns can be discerned from the male data (Table 3). For example, camping by trailer, camper or motor home is a popular activity of those males with some or no high school education and by those who have completed some or all of a college education. The \$4,000-\$5,999 and \$8,000-\$9,999 income groups are also disproportionately represented among the participants in this activity, as are those under 25 years of age, between 35 and 44 and over 55 years of age. This would appear to have some interesting sociological implications for campground operators.

### Satisfactions from Leisure Time Pursuits

The relative importance of satisfactions (irrespective of particular pursuits) is described in Table 5. Table 5 presents the averages or mean rating on a 1-5 scale of importance (1 = "very important to me"; 5 = "not important at all to me"), therefore, the lower the average, the more important that satisfaction is to most people across all the activities selected as "most favorite."<sup>5</sup>

Among women consumers, the most important satisfactions are "peace of mind," "chance to learn about new things," and "chance to get the most out of life while I can still enjoy it." A chance to "escape home or family pressures" is also important to women.

Among men consumers, the most important satisfactions are "peace of mind," "chance to get the most out of life while I can still enjoy it," and "adventure and excitement." The comfortableness of "an old familiar activity," and the generation of "happy memories" after the occasion has passed are also important satisfactions to men.

### Related Leisure Time Pursuits

The leisure time pursuits of consumers were examined to determine if clusters of related pursuits exist. Table 6 discloses the results of this analysis. This is a test of Hypothesis 2 and the results indicate that leisure time pursuits do cluster together.

The methodology employed in Table 6 is a factor analysis of the 50 leisure time pursuits indicated by respondents as those in which they engaged during the previous year. The analytical approach used was principal components factor analysis followed by Varimax rotation to simple structure. Groups were determined on the basis of statements which loaded at least 0.40 on the particular rotated factor and were at least three in number per rotated factor. Factors were extracted as long as the Varimax rotated factors accounted for at least 5 percent of variance, and had at least three variables loading 0.40 or greater.

Female leisure pursuits. Group 1 of female leisure pursuits (Table 6) appears to be an active, group sports oriented cluster while Group 2 is more individualistic in orientation. Group 3 represents a variety of pursuits in which the nature of the relationship is unclear. Group 4 seems to be active, people-oriented while group 5 appears individualist in orientation but also of a passive orientation. Groups 6 through 9 are definitely outdoor-oriented, but variously more active or more passive in orientation. These groups account for 51.2 percent of the variance.

Male leisure pursuits. Group 1 of male leisure pursuits (Table 6) appears to be active and traditionally oriented. Group 2 seems to be creative, passive, and perhaps delicate in orientation. Group 3 appears to be an upper socioeconomic scale orientation, while group 4 may be lower socioeconomic scale and people-oriented. Group 5 seems individualistic and outdoor while group 6 appears more passive and "around-the-house" in orientation. These groups account for 42 percent of the variance.

On the basis of these results, the first part of the hypothesis (H<sub>2</sub>) is accepted that leisure time pursuits can be clustered together in identifiable groups of activities.

TABLE 5

Average Rating of Each "Satisfaction"  
Statement Across All of the "Most  
Favorite" Activities

---

<u>Satisfaction Statement</u>	MALES (N=501)	FEMALES (N=589)
I feel that I am being creative	2.76	2.49
It gives me a chance to meet new people	2.63	2.61
It gives me a chance to learn about new things	2.24	1.94
I like it because it brings me into contact with friends	2.23	2.26
It provides me with a mental challenge: a problem to solve	2.41	2.38
It brings me peace of mind	1.87	1.75
It gives me a chance to experiment my style of living	2.62	2.59
It provides me with an escape from home or family pressures	2.29	2.06
It gives me a chance to develop a skill	2.22	2.19
It brings our family closer together; it helps achieve stronger family ties	2.31	2.32
It gives me a feeling of independence and self-reliance	2.29	2.25
It gives me a chance to be alone with my thoughts	2.41	2.27
I like it because it is an old familiar activity one with which I have had lots of experience	2.15	2.18
I like it because there is adventure and excitement in it	2.00	2.21
It provides an educational experience for my children	2.71	2.69
I like it because I have a feeling of mastery of the activity	2.28	2.31
It gives me a chance to get the most out of life while I can still enjoy it	1.91	2.00
It provides interesting experiences which I can tell my friends about afterwards	2.26	2.48
It gives me a chance to compete with others	2.61	3.07
I like it because of the uncertainty involved; a lot of different unexpected things can happen	2.33	2.66
It helps to keep me healthy and should prolong my life	2.32	2.54
I like it because I like to do things that will be of benefit to society or the community	3.00	2.92
It helps me to understand myself better	2.59	2.45
It provides me with a physical challenge, or a chance for intense physical activity	2.39	2.62
I like it because it brings happy memories to mind after the occasion has passed	2.17	2.12
It gives me an opportunity to seek out and enjoy the wonders of nature	2.43	2.58
I like it because it gives me a feeling of complete control over the outcome of the activity - what happens is strictly up to me	2.35	2.42
I feel I can respect myself for doing these things	2.28	2.19
It gives me an opportunity to see and do new and different things	2.38	2.20
It gives me a chance to be alone in a quiet, peaceful spot	2.44	2.29
It brings me recognition from other people	2.84	2.97
I like it because it helps me in my work	3.01	3.01

TABLE 6  
Groups (Factors) of Related Leisure-Time  
Pursuits Engaged in During 1972 by Females and Males

Female Respondents<sup>a</sup>

1. (7.2%)<sup>c</sup>  
Tennis (.780)<sup>d</sup>  
Bicycling (.653)  
Bowling (.518)  
Playing Basketball, Football, Baseball,  
Softball, Volleyball, Handball (.498)  
Chess, Checkers, Backgammon (.438)
2. (6.8%)  
Reading the Bible (.714)  
Gardening, Lawn Care, Landscaping (.648)  
Exercising, Jogging, Visiting Health Spa (.497)  
Church-Related Activities (.434)  
Horseback Riding (.427)
3. (6.0%)  
Volunteer, Community, School, Youth Group,  
or Charitable Organization Work (.718)  
Woodworking, Metalworking, Furniture Refinishing,  
Home Workshop Projects (.713)  
Not Horseback Riding (.535)
4. (6.0%)  
Square-Dancing or other organized dances (.703)  
Visiting a bar or club (.651)  
Photography, Taking pictures (.604)  
Playing piano, organ or other musical  
instrument for pleasure (.502)
5. (5.5%)  
Bingo, bridge or similar card games (.584)  
Attending sporting events as a spectator (.558)  
Collecting coins, stamps, bottles, etc. (.539)  
Swimming (.401)
6. (5.4%)  
Camping by Tent (.817)  
Attending Concerts or Plays (.552)  
Camping by trailer, camper, or motor home (.409)
7. (5.0%)  
Ice skating, Roller skating (.672)  
Canoeing, Rowing, Rafting (.605)  
Driving around for pleasure or sightseeing (.511)  
Playing Basketball, Football, Baseball, Softball,  
Volleyball, Handball (.403)
8. (4.7%)  
Hiking, Backpacking, Nature Study (.754)  
Walking for Pleasure (.515)  
Canoeing, Rowing, Rafting (.481)
9. (4.6%)  
Power boating, water skiing, scuba diving (.739)  
Fishing or Hunting (.696)  
Camping by trailer, camper, or motor home (.518)

Male Respondents<sup>b</sup>

1. (10.7%)<sup>c</sup>  
Attending sporting events as a spectator (.682)<sup>d</sup>  
Playing basketball, football, baseball, etc. (.627)  
Automobile modification or tune-ups (.607)  
Swimming (.598)  
Chess, checkers, backgammon (.590)  
Bingo, bridge or similar card games (.522)  
Bowling (.515)  
Pool, billiards or table-tennis (.446)  
Fishing or hunting (.401)
2. (8.2%)  
Playing the piano, organ or other musical instrument  
for pleasure (.657)  
Reading a book for pleasure (.654)  
Creative crafts or handicrafts (.563)  
Walking for pleasure (.530)  
Playing with children (.525)  
Woodworking, Metalworking, Furniture Refinishing, Home  
Workshop Projects (.516)  
Golf (.455)  
Listening to music from records, tape or radio (.453)  
Photography, Taking pictures (.421)
3. (6.4%)  
Hiking, Backpacking, Nature Study (.687)  
Attending concerts or plays (.662)  
Volunteer, Community, School, Youth Group, or  
Charitable Organization Work (.601)  
Golf (.603)  
Job-related reading or study (.496)  
Power boating, water skiing, scuba diving (.463)
4. (6.0%)  
Square-dancing or other organized dances (.758)  
Horseback riding (.730)  
Visiting a bar or club (.566)  
Ice skating, roller skating (.452)  
Bicycling (.444)
5. (5.7%)  
Canoeing, Rowing, Rafting (.819)  
Writing letters, doing crossword puzzles (.626)  
Camping by tent (.560)
6. (4.7%)  
Fixing up the house, remodeling, making repairs (.744)  
Gardening, Lawn Care, Landscaping (.606)  
Collecting coins, stamps, bottles, etc. (.470)

Notes:

- <sup>a</sup>Number of respondents indicating that they had engaged in a particular activity during 1972 varied from 63 to 507.
- <sup>b</sup>Number of respondents indicated that they had engaged in a particular activity during 1972 varied from 78 to 378.
- <sup>c</sup>Percentage of variance in the data accounted for by the total Varimax rotated factor.
- <sup>d</sup>Number in parentheses after leisure-time pursuit is the loading of that pursuit on that factor.



### Related Leisure Time Satisfaction

Satisfactions from leisure time pursuits were examined and were also found to cluster in identifiable groups. These groups are presented in Table 7. Similar methods of factor analysis were used in Table 7 as in the preceding analysis of leisure time pursuits. Statements of respondents refer to the satisfactions derived from their "most favorite" activities.

Female satisfactions from leisure time activities. Group 1 satisfactions for females (Table 7) stresses "newness" and "relating to people." Group 2 stresses mental activity and psychological independence, control and mastery. Group 3 has an active, physical, "body" orientation while group 4 is more contemplative and passive. Group 5 seems to imply a seeking of the unknown and of overcoming challenges. Group 6 may be an introspection-through-extroversion orientation, or as it is sometimes termed in clinical psychology, a "messiah complex." Group 7 implies a concern with self-respect. These groups account for 53 percent of the total variance.

Male satisfactions from leisure-time activities. Satisfactions from leisure time activities do not group together as clearly for males as for females. Within the first five groups, six statements appear in more than one group which implies less independent clusters of satisfactions.

Group 1 of male satisfactions (Table 7) has an implication of active, physical, competence-seeking. Group 2 seems to represent a desire for recognition through "good works," as well as seeking an understanding of self--perhaps the "messiah complex" again. Group 3 implies contemplative introspection, possibly outdoor-oriented. Group 4 may have a nostalgia orientation along with "living life to the fullest." Group 5 appears to represent a seeking of the unknown and the overcoming of challenges. Group 6 may represent a family orientation, while group 7 is oriented toward interaction with people. These groups account for 53 percent of the total variation.

On the basis of these results, the second half of the hypothesis ( $H_2$ ), that leisure time satisfactions can be clustered together in identifiable groups of satisfactions, is accepted.

### Related Satisfactions and Leisure Time Pursuits

Leisure time satisfactions and leisure time pursuits were analyzed to test the relationships between them. The results of this analysis are displayed descriptively in Tables 8 (females) and 9 (males). As in Table 5, the cell values in the matrix are the mean of the responses to each satisfaction statement for each of the listed pursuits. The lower the cell value, the more respondents felt that a particular satisfaction was an important outcome of their participation in the pursuit. The full description of the volume of data in these two tables precludes a detailed verbal analysis in this paper; therefore, only a few selected relationships will be mentioned.

Among the female respondents, the strongest benefits or satisfactions derived from camping by trailer, camper, or motor home were "stronger family ties" and "enjoying the wonder of nature." Similarly, gardening, lawn care or landscaping presents a good opportunity to be "alone with my thoughts" for many of the women. This particular cell value is also the lowest (1.09).

TABLE 7

Groupings (Factors) of Related Satisfaction Statements  
Across All Activities in "Most Favorite" Category For Females and Males

<u>Female Respondents<sup>a</sup></u>	<u>Male Respondents<sup>b</sup></u>
<p style="text-align: center;"><u>1. (9.5%)<sup>c</sup></u> Meet new people<sup>d</sup> (.653)<sup>e</sup> Learn new things (.621) Contact with friends (.603) Do new things (.589) Happy memories (.533) Stronger family ties (.517) Get most out of life (.486)</p> <p style="text-align: center;"><u>2. (9.1%)</u> Develop a skill (.791) Being creative (.721) Feeling of mastery (.606) Mental challenge (.563) Feeling of independence (.535) Feeling of control (.400)</p> <p style="text-align: center;"><u>3. (8.3%)</u> Physical challenge (.520) Keep healthy (.758) Enjoy wonders of nature (.612) Stronger family ties (.463)</p> <p style="text-align: center;"><u>4. (8.0%)</u> Alone in quiet spot (.797) Alonewith thoughts (.743) Peace of mind (.548)</p> <p style="text-align: center;"><u>5. (6.6%)</u> Uncertainty involved (.660) Interesting experiences (.626) Chance to compete (.610) Adventure and excitement (.524) Recognition from others (.426)</p> <p style="text-align: center;"><u>6. (6.1%)</u> Benefit to society (.727) Helps in my work (.705) Understand myself better (.503) Educational for children (.417) Recognition from others (.400)</p> <p style="text-align: center;"><u>7. (5.2%)</u> Can respect myself (.749) Feeling of control (.566) Get most out of life (.413)</p>	<p style="text-align: center;"><u>1. (9.3%)<sup>c</sup></u> Physical challenge<sup>d</sup> (.712)<sup>e</sup> Keep healthy (.698) Develop a skill (.664) Feeling of mastery (.625) Chance to compete (.490) Feeling of control (.482) Feeling of independence (.403)</p> <p style="text-align: center;"><u>2. (9.1%)</u> Helps in my work (.697) Recognition from others (.639) Being creative (.600) Benefit to society (.589) Understand myself better (.535) Mental challenge (.511) Learn new things (.435)</p> <p style="text-align: center;"><u>3. (8.2%)</u> Alone in quiet spot (.784) Alone with thoughts (.770) Enjoy wonders of nature (.512) Peace of mind (.474) Feeling of independence (.454) Do new things (.414) Style of living (.411)</p> <p style="text-align: center;"><u>4. (8.2%)</u> Happy memories (.655) Old familar activity (.655) Get most out of life (.548) Can respect myself (.543) Interesting experiences (.456) Peace of mind (.455) Feeling of control (.451)</p> <p style="text-align: center;"><u>5. (6.8%)</u> Uncertainty involved (.680) Adventure and excitement (.577) Chance to compete (.565) Interesting experiences (.504) Do new things (.428)</p> <p style="text-align: center;"><u>6. (5.9%)</u> Stronger family ties (.671) Educational for children (.661) Benefit to society (.490) Understand myself better (.421) Enjoy wonders of nature (.405)</p> <p style="text-align: center;"><u>7. (5.7%)</u> Meet new people (.765) Contact with friends (.754) Learn new things (.530)</p>

## Notes:

<sup>a</sup>Number of respondents answering these questions varied from 580 to 589.

<sup>b</sup>Number of respondents answering these questions varied from 489 to 502.

<sup>c</sup>Percentage of variance in the data accounted for by the total Varimax rotated factor.

<sup>d</sup>Paraphrased satisfaction statements. Refer to Table 5 for complete statements.

<sup>e</sup>Loading of that statement on the factor.

TABLE 8  
Mean Values of "Satisfactions" Statements for the  
16 Most Popular Favorite Activities  
(Female)

	Being creative <sup>c</sup>	Meet new people	Learn new things	Contact with friends	Mental challenge	Peace of mind	Experiment with style of living	Escape from pressures	Develop a skill	Develop stronger family ties	Independence and self-reliance	Alone with my thoughts	Old familiar activity	Adventures and excitement	Educational for the kids	Feeling of mastery	Get the most out of life	Interesting experiences	Chance to compete	Uncertainty involved	Keeps me healthy	Benefits the community or society	Understand myself better	Physical challenge	Brings happy memories	Enjoy wonders of nature	Control over the outcome	Can respect myself	See and do new things	Alone in a quiet spot	Recognition from others	Helps me in my work
LEISURE-TIME PURSUITS																																
Attending movies (45) <sup>b</sup>	2.78	2.96	2.13	2.37	3.10	2.21	3.13	1.66	3.11	2.26	2.89	2.95	2.71	2.34	2.48	3.07	2.04	2.36	3.32	2.76	3.22	3.22	2.73	2.22	2.22	2.91	2.71	2.70	2.35	2.72	3.24	3.15
Bingo, bridge, etc. (77)	2.94	2.30	2.74	1.79	1.82	2.43	3.06	2.27	2.27	2.57	2.74	3.10	2.18	2.54	2.98	2.32	2.40	2.73	2.00	2.25	3.09	3.07	2.73	2.46	2.46	3.24	2.78	2.61	2.99	3.18	3.10	3.17
Bowling (52)	3.01	1.81	2.66	1.36	2.45	2.42	2.95	2.74	1.91	2.30	2.52	2.25	2.51	2.22	3.13	2.51	2.04	2.36	1.86	2.24	1.89	3.27	1.99	1.95	1.95	2.80	2.91	1.87	2.36	2.91	3.11	2.74
Church-related activities (44)	2.43	1.78	1.87	1.73	2.28	1.78	2.85	2.68	2.73	1.82	2.69	2.72	2.23	2.37	2.32	2.74	2.06	2.38	3.21	2.74	2.67	1.93	1.99	1.95	1.95	2.80	2.91	1.87	2.36	2.91	3.11	2.74
Gardening, lawn care, etc. (66)	1.88	3.08	2.07	2.89	2.46	1.46	2.20	2.21	2.04	2.27	2.15	1.09	2.31	2.75	2.58	2.08	2.02	2.88	3.15	2.55	2.08	2.64	2.62	2.45	2.45	1.59	2.23	1.78	2.40	1.92	2.97	3.07
Concerts or plays (44)	3.06	2.44	2.09	2.19	2.96	2.29	2.96	2.05	3.19	2.50	3.07	3.08	2.60	2.08	2.40	3.14	2.07	2.23	3.34	2.94	3.28	2.82	2.63	1.90	1.90	3.20	3.33	2.89	2.01	3.25	3.34	3.30
Camping by trailer, etc. (65)	3.02	1.82	1.95	1.78	3.03	1.52	2.09	1.63	2.07	1.28	2.39	2.37	2.51	1.50	1.79	2.96	1.69	1.97	3.72	2.51	2.10	3.04	2.49	1.49	1.49	1.27	2.90	2.29	1.40	1.79	3.51	3.41
Fishing or hunting (77)	3.03	2.88	2.54	2.40	2.52	1.67	2.76	1.68	2.17	1.66	2.34	1.83	2.04	1.80	2.47	2.45	1.95	1.89	2.88	2.49	2.35	3.20	2.73	1.77	1.77	1.68	2.61	2.46	2.42	1.91	3.38	3.18
Picnicking (62)	2.90	2.79	2.73	2.07	3.23	1.95	2.71	1.98	3.17	1.44	2.82	3.17	1.93	2.35	2.38	2.89	1.88	2.50	3.47	3.01	2.69	3.01	2.95	1.99	1.99	1.83	3.01	2.54	2.37	2.55	3.52	3.32
Swimming (70)	3.11	3.17	2.79	2.54	2.95	2.14	2.83	1.86	2.21	2.31	2.04	2.33	2.02	2.41	2.41	2.25	1.99	3.08	3.07	2.48	1.62	3.21	2.76	2.30	2.30	2.61	2.54	2.38	2.77	2.27	3.14	3.19
Creative crafts or handicrafts (257)	1.32	3.01	1.72	2.82	1.94	1.80	2.47	2.27	1.42	2.80	2.05	1.87	2.14	2.79	2.84	1.89	2.20	2.69	3.08	2.85	3.05	2.90	2.56	2.55	2.55	3.12	1.92	1.84	2.15	2.15	2.82	2.84
Driving around for pleasure (52)	3.27	2.17	1.56	2.30	3.03	1.76	2.88	1.91	2.95	1.79	2.48	2.16	2.35	1.92	2.22	2.92	1.85	2.10	3.27	2.28	2.82	2.84	2.68	1.77	1.77	1.58	2.77	2.50	1.61	2.25	3.00	3.04
Listening to music (54)	3.02	3.16	2.89	3.04	2.89	1.41	2.73	1.98	3.18	2.72	2.84	1.72	2.03	3.08	3.07	2.97	2.27	2.90	3.46	3.18	3.14	3.22	2.48	2.26	2.26	3.08	3.04	2.44	2.72	1.88	3.12	2.91
Playing with children (46)	1.81	2.44	1.85	2.09	1.86	1.75	2.29	2.70	2.06	1.64	2.43	3.05	2.00	1.83	1.97	2.72	1.84	2.09	3.48	2.12	2.30	2.38	1.94	1.84	1.84	2.44	3.01	1.93	2.02	3.21	3.05	2.57
Reading a book for pleasure (154)	3.06	3.44	1.52	3.31	2.38	1.71	3.10	1.94	2.81	3.05	2.82	2.06	2.19	1.98	2.83	2.66	2.43	2.31	3.03	2.89	3.16	3.07	2.13	2.40	2.40	2.86	3.06	2.35	2.61	1.62	3.33	2.64
Visiting with friends, partying (94)	2.87	1.61	1.88	1.27	3.00	2.48	2.60	2.15	3.24	2.34	2.79	3.32	2.54	2.50	2.92	2.92	1.74	2.56	3.21	2.58	2.94	2.94	2.31	1.61	1.61	2.64	2.93	2.52	2.07	3.38	3.05	3.05

Notes: Recorded value is the mean of the summed 3 categories of "most favorite," "second most favorite" and "third most favorite" activity. There were no significantly bimodal distributions on any of the satisfaction statements - distributions tended to be either skewed, peaked, or approximately "normal."

<sup>b</sup>Number in parentheses is number of participants.

<sup>c</sup>Paraphrased "Satisfactions" Statement

TABLE 9  
Mean Values of "Satisfactions" Statements for the 15 Most Popular Favorite Activities<sup>a</sup>. (Male)

LEISURE-TIME PURSUITS	Being creative <sup>c</sup>	Meet new people	Learn new things	Contact with friends	Mental challenge	Peace of mind	Experiment with style of living	Escape from pressures	Develop a skill	Develop stronger family ties	Independence and self-reliance	Alone with my thoughts	Old familiar activity	Adventures and excitement	Educational for the kids	Feeling of mastery of life	Interesting experiences	Chance to compete	Uncertainty involved	Keeps me healthy	Benefits the community or society	Understand myself better	Physical challenge	Brings happy memories	Enjoy wonders of nature	Control over the outcome	Can respect myself	See and do new things	Alone in a quiet spot	Recognition from others	Helps me in my work
Attending sporting events (54) <sup>b</sup>	3.34	2.55	2.50	2.01	2.85	2.51	3.11	2.27	3.16	2.38	2.99	3.29	2.45	2.05	2.78	3.02	2.27	2.38	2.97	2.98	3.01	3.01	2.92	2.36	3.07	3.24	2.92	2.70	3.28	3.19	3.05
Playing basketball, football, etc. (45)	2.78	2.19	2.51	1.58	2.48	2.19	2.62	2.03	1.71	2.71	2.24	2.90	2.04	1.58	2.83	1.90	1.63	2.38	1.54	1.61	2.96	2.84	1.43	2.18	3.11	2.39	2.20	2.60	3.05	2.46	3.12
Bowling (49)	3.03	2.05	2.49	1.69	2.12	2.25	2.97	2.31	1.82	2.58	2.60	3.21	2.07	2.13	2.97	2.05	2.14	2.33	1.68	2.04	3.23	2.88	1.85	2.23	3.31	2.28	2.34	2.98	3.21	2.42	3.15
Fixing up the house (42)	1.68	3.13	1.79	3.22	1.67	2.01	2.08	2.51	1.93	2.26	1.96	2.42	2.24	2.62	2.54	2.15	2.10	2.46	2.90	2.48	2.72	2.39	2.34	2.19	2.94	2.00	2.09	1.78	2.51	2.69	2.51
Gardening, lawn care, etc. (71)	1.83	3.24	2.65	2.98	2.50	1.33	2.56	2.64	2.33	2.38	2.06	2.24	2.35	2.75	2.98	2.36	2.23	2.78	3.07	2.07	2.92	2.79	2.22	2.66	1.99	2.53	2.28	2.55	2.28	2.53	3.01
Camping by trailer, etc. (36)	2.92	2.27	2.30	2.11	2.95	1.62	2.38	2.18	2.69	1.45	2.39	2.48	1.70	2.03	2.20	2.79	1.77	2.30	3.61	2.08	3.11	2.84	2.41	1.61	1.51	2.61	2.24	1.77	2.22	3.29	2.81
Fishing or hunting (64)	3.10	2.81	2.40	2.32	2.58	1.58	2.57	2.00	2.17	2.12	2.00	1.83	1.82	1.65	2.46	2.19	1.81	2.07	2.52	2.03	3.18	2.65	2.18	1.95	1.65	2.43	2.26	2.26	1.71	3.04	3.17
Golf (44)	3.49	2.48	2.98	1.94	2.14	1.38	3.00	2.15	1.85	3.15	2.30	2.35	2.49	2.43	3.18	2.18	1.92	2.32	1.82	1.82	3.33	2.66	1.98	2.31	2.84	1.91	2.28	2.74	2.55	2.80	3.36
Power boating, water skiing, etc. (40)	3.30	2.30	2.29	2.06	2.87	1.90	2.66	2.38	2.14	1.73	2.05	2.57	2.34	1.46	2.59	2.06	1.60	2.58	2.69	2.01	2.25	3.09	2.14	1.90	1.61	1.89	1.93	1.92	2.73	3.22	3.55
Swimming (49)	3.36	2.87	3.36	2.35	3.02	1.89	2.84	2.18	2.15	2.27	2.45	2.49	2.10	2.46	2.74	2.14	1.85	2.83	2.93	1.67	3.18	2.82	1.61	2.44	2.43	2.05	3.07	2.75	3.15	3.30	
Driving around for pleasure (49)	3.09	2.87	1.90	2.34	3.16	1.76	3.08	2.20	3.15	1.72	2.46	2.74	2.18	2.13	3.32	2.67	1.77	2.25	3.30	2.77	2.19	2.71	2.97	1.73	1.56	2.56	2.62	1.58	2.45	3.21	3.08
Listening to music (60)	3.07	3.41	2.61	3.20	3.43	1.56	2.89	2.44	3.09	2.68	3.14	1.86	2.59	2.91	3.05	3.30	2.37	3.07	3.62	3.24	3.38	2.85	3.49	2.42	3.13	2.91	2.42	3.09	2.03	3.34	3.12
Reading a book for pleasure (47)	3.10	3.64	1.38	3.56	2.17	1.48	2.71	2.44	2.73	3.02	2.68	1.62	1.96	2.23	2.81	2.85	2.17	2.49	3.38	3.23	3.29	2.31	3.33	2.45	2.77	2.88	2.27	2.50	1.73	3.28	2.61
Visiting with friends, partying (49)	2.98	1.96	2.14	1.29	2.41	2.41	2.70	2.34	2.89	2.10	2.74	2.65	2.72	2.36	3.02	3.14	1.88	2.31	3.14	2.95	2.82	2.41	2.83	1.98	2.66	2.71	2.50	1.98	2.99	2.94	2.97
Woodworking, metal working, etc. (50)	1.49	3.14	2.05	2.90	1.74	1.95	2.45	2.45	2.09	2.77	2.01	2.23	2.09	2.69	2.63	2.15	2.20	2.73	3.02	2.66	2.79	2.62	2.43	2.61	2.98	1.89	2.18	2.26	2.36	2.76	2.70

Notes: <sup>a</sup>Recorded value is the mean of the summed 3 categories of "most favorite," "second most favorite" and "third most favorite" activity. There were no significantly bimodal distributions on any of the satisfaction statements - distribution tended to be either skewed, peaked, or approximately "normal."

<sup>b</sup>Number in parentheses is number of participants.

<sup>c</sup>Paraphrased "Satisfactions" Statement

Among the male respondents, for example, the strongest satisfactions from playing basketball, football, baseball, softball, volleyball or handball, are "contact with friends," a "chance to compete," and a "physical challenge." The men indicated that reading a book for pleasure provided the greatest opportunity to be "alone with my thoughts," rather than gardening, lawn care, etc. as the women had indicated. The lowest cell value in this table (1.29) indicates that "contact with friends" is the strongest benefit to men from visiting with friends or partying.

### Attitudes Toward Leisure Time Activities

The attitudes toward leisure time activities of consumers were measured by the use of AIO (activity, interest, opinion) statements. Some of the more illuminating responses to activity statements are presented in Table 10.

Vacations. An inspection of activity statements relating to vacations (Table 10) discloses that Americans are family-oriented in their vacations. Both males and females agree that their family often travels together. There is also substantially the same proportion of responses from males and females concerning the type of vacation that is appropriate, and the importance of the children in vacation planning.

Entertainment. Television is the primary source of entertainment for a family and both males and females agree about this (Table 10). There is substantial agreement about entertainment preferences although females have somewhat more preference for an evening out (dinner or theatre) and males have slightly stronger preference for a quiet evening at home. Sporting events are rated considerably higher by men than by women as a leisure activity, with men putting somewhat more emphasis on the competitive value on sports and on participation.

Value of time. More Americans believe they do not have enough leisure time than believe they do have enough time (Table 10). This is especially true for men compared to women. Most people are relatively satisfied with their leisure activities and do not consider them boring and, therefore, apparently wish they had more time for those activities. To a large degree, consumers appear to rate the factor of time as more important than money in their recreation (52% of women, 55% of men).

### Conclusions

This study, while exploratory in nature, offers encouragement to both the theoretician of leisure and consumer choice, and to the business strategist interested in the profit opportunities of time and leisure markets.

First, this study indicates it is possible to measure the satisfactions that consumers derive from leisure time activities and in this sense is an expansion of the work of Donald and Havighurst (1959). Out of the 32 "satisfaction" statements, 21 of them appeared to be meaningful to the female respondents in relation to pursuits selected as favorites. Twenty-three of these statements appeared meaningful to males.

It should be noted that individuals can apparently relate their participation in outdoor, active, group-oriented pursuits more readily to derived

TABLE 10

## Selected Statements Related to Individual and Family Activities by Females and Males

Activity Statements	FEMALES (N=594)					MALES (N=490)				
	SA	A	N/O	D	CD	SA	A	N/O	D	CD
<b>Vacation Related</b>										
Our family travels together quite a lot . . . . .	38%	30%	7%	16%	9%	37%	33%	7%	15%	8%
A cabin by a quiet lake is a great place to spend the summer . . . . .	44	28	10	11	7	45	30	9	11	5
On a vacation, I just want to rest and relax . . . . .	31	30	7	23	9	34	30	4	21	11
I like to spend my vacations in or near a big city . . . . .	6	13	12	30	39	4	10	11	28	47
On my vacations, I like to get away from mechanization and automation . . . . .	23	33	16	19	9	28	37	14	16	5
Vacations should be planned for children . . . . .	17	39	16	20	8	18	38	19	18	7
<b>Entertainment Related</b>										
Television is our primary source of entertainment . . . . .	26%	26%	6%	23%	19%	24%	31%	6%	19%	20%
I would rather spend a quiet evening at home than go out to a party . . . . .	24	30	8	25	13	30	31	8	23	8
We do not often go out to dinner or the theater together . . . . .	20	22	8	19	31	15	25	6	22	32
<b>Sporting Related</b>										
The best sports are very competitive . . . . .	13%	21%	31%	21%	14%	28%	28%	20%	15%	9%
I prefer to participate in individual sports more than team sports . . . . .	11	18	39	18	14	16	29	25	17	13
Whenever possible, I prefer to participate in sporting activities, rather than just watch them . . . . .	15	27	15	19	24	25	29	12	18	16
I like to go and watch sporting events . . . . .	18	40	13	15	14	34	37	10	12	7
<b>Leisure Time Related</b>										
I have enough leisure-time . . . . .	14%	23%	8%	28%	27%	13%	14%	8%	31%	34%
I tend to spend most of my leisure-time indoors . . . . .	16	35	6	29	14	7	22	7	32	32
Basically, I'm satisfied with my present leisure-time activities . . . . .	21	45	7	20	7	25	39	7	22	7
My leisure-time tends to be boring . . . . .	4	14	8	26	48	4	12	8	27	49
<b>Specific Activity Related</b>										
I do a lot of repair work on my car . . . . .	1%	4%	23%	7%	65%	24%	24%	5%	16%	31%
I often work on a do-it-yourself project in my home . . . . .	37	34	15	7	7	36	31	13	11	9
I am active in one or more service organizations . . . . .	12	11	19	17	41	8	10	23	18	41
<b>General Statements</b>										
When it comes to my recreation, time is a more important factor to me than money . . . . .	23%	29%	17%	21%	10%	25%	30%	15%	19%	11%
When it comes to my recreation, money is a more important factor to me than time . . . . .	9	19	16	34	22	10	20	18	34	18
I watch television more than I should . . . . .	21	28	7	23	21	20	29	9	24	18
My major hobby is my family . . . . .	49	30	9	9	3	35	32	14	14	5

## Note:

SA - Strongly Agree; A - Agree Somewhat; N/O - Undecided or no opinion;  
D - Disagree Somewhat; CD - Completely Disagree

satisfactions than they can participation in indoor, passive, primarily individual pursuits. The latter type of pursuit engendered an "indifferent or does not apply" response on half or more of the satisfactions statements by half or more of the respondents. This might imply that people who participate in this type of pursuit do so for reasons (to gain "satisfactions") which are (1) subconscious, (2) perceived to be socially unacceptable in some sense, (3) not included in the list provided, or (4) few in number compared with outdoor, active, group-oriented pursuits.

Second, this study indicates that leisure time pursuits do cluster together in interpretable groups (factors) which individually account for approximately 5-16 percent of data variance. The clustering of leisure time pursuits supports the earlier findings of Proctor (1962), Burton (1971), and others.

A distinct difference was found between the indoor orientation of females and the slight overall outdoor orientation of the males. There were 13 indoor-oriented pursuits selected as favorites by females, but only eight outdoor-oriented ones. The males, on the other hand, selected 13 indoor-oriented pursuits as favorites (not the same set as the females) and 15 outdoor-oriented ones. In addition to the generally narrower range of pursuits selected by women, it is also significant to note that females and males appear to have sufficiently diverse favorites in some cases as to be incompatible in terms of simultaneous performance. The resolution of this incompatibility should be of interest to family sociologists as well as to marketing strategists.

Leisure time satisfactions group together across many pursuits and also in relation to a single pursuit. This is significant because it indicates that individuals seek out a group of related leisure time satisfactions through a variety of means. Women appear to seek people contact, novelty, memories and stronger family relations in their favorite activity, whereas men seek challenge, mastery, control, recognition, and independence.

Certain pursuits appear to perform a "linking-pin" function in connecting two or more groups of (internally) "strongly" related pursuits. Picnicking and photography are examples of such pursuits. These linking pursuits appear to provide bundles of satisfactions which are different enough to satisfy different groups of participants.

The data in this study indicate many possibilities for determining segments of the markets on the basis of the satisfactions desired rather than the pursuits. This opens the possibility--both for planners of public facilities as well as strategists in the private sector--of substituting one leisure pursuit of lower cost or different availability for pursuits which have become difficult to make available. Conversely, leisure strategists may wish to conduct additional research along these lines in order to plan the grouping of recreation facilities in such a way as to offer related satisfactions to participants in a particular market segment (refer to National Academy of Sciences, 1969). Another application might be the planning of leisure facilities to provide alternative satisfactions for various family members within a close physical proximity.

## FOOTNOTES

1. This research was supported by the Fred and Mabel Dean Hill Research Fund. The authors also express appreciation to Professor Arthur Cullman, The Ohio State University, for his encouragement and assistance in the completion of this project. The senior author also wishes to express his appreciation to Ms. Tammy Tulley and Mr. Staale Engen for their able assistance in the computations resulting in Tables 2, 3, 8, and 9.
2. Douglass K. Hawes is Assistant Professor of Business Administration, University of Wyoming.
3. W. Wayne Talarzyk is Associate Professor of Marketing, The Ohio State University.
4. Roger D. Blackwell is Professor of Marketing, The Ohio State University, and Vice President, Management Horizons, Inc.
5. Respondents were asked to select their "most favorite," "second most favorite" and "third most favorite" pursuits from the list of 50 pursuits provided. Each of these three pursuits were then rated against each of the 32 "satisfactions" statements. Table 5 presents the averages of just the "most favorite" category.

## REFERENCES

- Bultena, G. L., & Klessig, L. L. Satisfaction in camping: A conceptualization and guide to social research. Journal of Leisure Research, 1969, 3, 348-354.
- Burton, T. L. Experiments in recreation research. London: George Allen and Unwin, Ltd., 1971.
- Donald, M. N., & Havighurst, R. J. The meanings of leisure. Social Forces, 1959, 37, 355-360.
- Ellis, M. J. Why people play. Englewood Cliffs: Prentice-Hall, 1973.
- Gray, D. E. Identification of user-groups in forest recreation and determination of the Characteristics of such groups, Unpublished doctoral dissertation, University of Southern California, 1961.
- Havighurst, R. J., & Feigenbaum, K. Leisure and life-style. American Journal of Sociology, 1959, 64, 396-404.
- Havighurst, R. J. Nature and values of meaningful free-time activity Robert W. Kleemier (Ed.), Aging and leisure. New York: Oxford, 1961.
- Hawes, D. K. An exploratory nationwide mail survey of married adult leisure-time behavior patterns and the satisfactions derived from leisure-time pursuits Unpublished doctoral dissertation, The Ohio State University, 1974.
- Kaplan, M. Leisure in America: A social inquiry. New York: Wiley, 1960.



- Kaplan, M. Leisure: Issues for American business. Paper presented at seminar on the economics of leisure, University of Denver, October 26, 1970.
- Kraus, R. Recreation and leisure in modern society. New York: Appleton-Century-Crofts, 1971.
- Linder, S. B. The harried leisure class. New York: Columbia University Press, 1970.
- Lowrey, G. A., Jr. A multivariate analysis of the relationship between selected leisure behavior variables and personal values. Unpublished doctoral dissertation, University of Illinois, 1969.
- Lundberg, G. A., Komarovsky, M., & McInerney, M. A. Leisure: A suburban study. New York: Agathon, 1969.
- Maw, R. Construction of a leisure model. Ekistics, 1971, 230-238.
- National Academy of Sciences. A Program for outdoor recreation research. Washington: National Academy of Sciences, 1969.
- Neulinger, J., & Breit, M. Attitude dimensions of leisure. Journal of Leisure Research, 1971, 1, 255-261.
- Proctor, C. H. Dependence of recreation participation on background characteristics of sample persons in the September 1960 national recreation survey. In Abbot L. Ferriss (Ed.), National recreation survey. Washington: Outdoor Recreation Resources Review Commission, Study Report 19, 1962, 77-94.
- Roberts, K. Leisure. London: Longman, 1970.
- Tatham, R. L., & Dornoff, R. J. Market segmentation for outdoor recreation. Journal of Leisure Research, 1971, 3, 5-16.
- Voss, J. The definition of leisure. Journal of Economic Issues, 1967, 1, 91-106.
- White, R. C. Social class differences in the uses of leisure. The American Journal of Sociology, 1955, 61, 145-150.

## MARKETS FOR LEISURE TIME

Justin L. Voss<sup>1</sup>  
State University of New York  
at Plattsburgh

Roger D. Blackwell<sup>2</sup>  
The Ohio State University

The authors examine the nature of consumers' resources and indicate how a redefinition of the time constraint provides theoretical insights into the consumer's market for leisure time. This paper indicated the emergence of a second market for leisure time. Historically, economists have identified the market for leisure time with a reduction in working hours. This practice evolved from mercantilist England, where workers could freely vary their hours of labor. In a modern economy, however, where leisure is no longer equated with non-work time, most consumers can buy time goods--goods and services which produce leisure time. Time goods constitute a second market for leisure time--a market analytically equivalent to a reduction in working hours. This paper suggests that this second market for leisure time represents a profitable opportunity for marketing strategies and a promising area for analytical and empirical research.

## Time and Money Budgets

The affluence of recent decades has led some consumer analysts and economists to reconsider the economic constraints on individual consumption. One result of this research has been the growing recognition that consumer resources consist of two budget constraints--a money budget and a time budget. Since most commodities require spending time as well as money, in a high-consumption economy scarcity of time becomes an increasingly important problem--possibly the major factor limiting many affluent consumers' economic freedom.

In spite of the importance of the problem of time-scarcity, current economic literature has progressed little beyond the stage of formulating an economic model which includes time and money as multiple budget constraints (Voss, 1964; Becker, 1965; Johnson, 1966; Linder, 1970). Little of this literature has given relevant business application. The literature of consumer research has been more adventuresome but less precise (Engel, Kollat, Blackwell, 1973, pp. 104-106). With some recent exceptions (Schary, 1971; Foote, 1966), market literature has treated lightly the analytical impact of time resources on consumer behavior. Adding a time dimension to consumer behavior, however, greatly affects consumer behavior and future consumption markets. Indeed, if consumer researchers are to penetrate the questions of time scarcity and respond effectively, they must develop perspectives on the scope and dimension of a problem that underlies many of the complexities of consumer behavior.

Since consumption of most goods and services requires spending time (which includes nearly all the leisure industry), in the future the issue facing both marketer and consumer will be much the same--one of discovering ways to augment individual time budgets so that consumption of preferred goods and services can be expanded and optimizing consumption patterns achieved. In a sense the future

of consumer welfare depends to a considerable extent on the answer to one critical question. How can consumers obtain more time? Or more specifically for consumer research, how and where can consumers buy more time.? The answer necessitates a thoroughgoing knowledge of the "The Market for Time."

Unfortunately, the literature of economics and consumer behavior has given little analytical recognition to the concept of a market for time, per se. Every day, however, consumers are buying discretionary or leisure time. Where is this market? What prices are being paid? What quantities are being purchased.? How efficient are the marketing institutions involved in these transactions? These questions must at least be raised and tentative answers given if consumer research is to meet the needs of future markets.

### Markets for Time

A clear conception of time prices and of leisure are essential to an analytical treatment of each of these concepts in consumer research.

#### Nature of the Time Constraint

In the conceptualization of many current authors, leisure is defined as:

$$L = T - W$$

where L = Leisure

T = Total time available

W = Work

By adopting the above conceptualization in the contemporary work-leisure model rather than a more sophisticated conceptualization, many insights into consumer behavior have been eliminated.

In this paper a more extended conceptualization of leisure is adopted. This conceptualization provides the basis for a more in-depth understanding of leisure-spending patterns and one that is more in accord with the consumer market behavior. The following conceptualization is used in this paper:

$$L = T - (N + W)$$

where L = Leisure

T = Total time

N = Nondiscretionary time

W = Work

Consumer research has quantified qualitative differences between distinct periods of time, which correspond to the division of nonwork time into nondiscretionary time and leisure. One researcher (Foote, 1966) has identified what he terms "special" activities as distinguished from "routine" activities. Significantly different preference ratings exist between the different

activities and Foote indicates that whereas most activities are considered "routine", "special" activities tend to be concentrated on weekends. This clustering of "special" or leisure time activities corresponds to the increased availability of time on weekends. A more thorough discussion of this definition of leisure and alternative conceptualizations of leisure is available in an earlier paper (Voss, 1964; 1967).

#### Analytical Models of Time and Leisure-- A Second Market for Time

Since the mid-1960s, a number of economists and consumer researchers have found weaknesses in contemporary explanations of the consumer's work-leisure behavior. Their arguments center on grounds other than the traditional arguments of imperfect markets or standard work weeks. For some reason, not readily apparent, these widely recognized arguments or criticisms of the contemporary model have not been included by modern critics. Certainly, it would have made their case stronger. In fact, Becker's alternative model is based on the foregone earnings concept and its corollary assumption of flexibility (Becker, 1965). These writers suggest, implicitly or explicitly, the need for a revision of the contemporary explanation for consumers' work-leisure behavior, even though they differ both as to why a reconstruction is necessary and the nature of the model.

They argue that a fundamental aspect of consumer behavior has been neglected. Namely, consumers spend from two resources: a time budget and a money budget. (Voss, 1964; Becker, 1965; Johnson, 1966; Linder, 1970; Schary, 1971).

#### The First Market for Time

In a contemporary model of consumer time behavior, leisure time,  $L$ , is defined as equal to total time available,  $T$ , minus work time,  $W$ . By definition, work and leisure are linked analytically. Ergo, leisure time is purchased only by reducing working hours. The price of leisure time is measured by foregone earnings. Proponents of this contemporary model argue that leisure as a consumer good differs only in the way it has to be bought. One gives up income, but it is done by working fewer hours at remunerative employment. The exchange of money for a good is still there by the way the transfer is accomplished (Brennan, 1965, p. 293).

Adopting a  $T-W$  definition of leisure time has distinct analytic advantages. It enables one model to explain both the demand for leisure time and the individual supply curve for labor. This dichotomization evolved from the classical work-leisure model where work was defined as pain, and leisure was defined as freedom from pain. The only way the individual could buy leisure (freedom from pain of labor) was to work less. Increasing leisure was tied by analytic necessity to decreasing working hours. Although historical working conditions have changed greatly, the contemporary work-leisure model adopts the  $T-W$  definition of leisure time along with its analytical corollaries--the model's duality and its specification of one market for time. Identifying and accepting changes in working hours as a market for time, however, is quite a different thing from accepting it as the market for time.

### The Second Market for Time

The position of the present paper is that the contemporary work-leisure model is deficient on definitional grounds. Since it is generally understood that when workers demand leisure time they are demanding the luxury of discretionary or "own" use of the time resource, it is apparent empirically that all nonwork time is not leisure. As Foote has demonstrated, consumers are willing to pay a price to reduce time expenditures on certain items in their activity set (Foote, 1966). Consequently, the contemporary model's definition of leisure,  $L = T - W$ , does not adequately characterize the leisure variable. All nonwork time simply is not leisure. This empirical observation has not been explained analytically in the context of the work-leisure model. When leisure is defined with greater empirical accuracy as  $L = T - (N + W)$ , it provides a more satisfactory explanation of consumer behavior and introduces the possibility of more than one market for time.

Once the dichotomization of work and leisure has been invalidated, there follow important analytical implications. If, after redefining leisure, the concept of foregone earnings is broadened to foregone income, then one may argue that income may be exchanged for leisure time directly, in the same manner that other commodities are exchanged without affecting working hours.  $L$  may be increased by reducing  $N$  while  $W$  remains unchanged. These nonwork reducing, leisure-time producing expenditures constitute a second market for time. Analytically, such consumer expenditures augment leisure time in the same sense as a reduction in working hours. These leisure time producing goods and services have been identified previously as time goods (Voss, 1967).

Moreover, when there are ways to buy leisure time other than through reducing working hours, the duality property of the contemporary work-leisure model is nullified. Indeed, while the demands for leisure clearly continue to provide a necessary and sufficient explanation for the individual's supply curve for labor in a modern economy, the reverse is not longer the case. A separate model is required to explain the demand for leisure time. The substitute relationship between work and leisure specified definitionally by the contemporary work-leisure model is replaced by a model capable of explaining both substitute and complementary relationships.

This second market for leisure time may be identified empirically. As early as 1932, at least one writer had identified this potential market and distinguished time-saving from labor-saving expenditures (Pitkin, 1932, pp. 231-232). Moreover, for centuries, domestic servants had served as a leisure-producing function even though their services were limited to a small segment of the consuming public.

Armed with the knowledge that consumers do, in fact, demand additional time to use at their own discretion, even the most casual observation of the present-day economic environment would reveal that consumers buy leisure, or "own" time, through the purchasing of goods and services with nondiscretionary time-saving, hence, leisure time augmenting characteristics.

From a marketing perspective, the reconstructed work-leisure model presented in this paper offers insight into the value of products. Servants have been apparently replaced by self-cleaning ovens, microwave ovens, and instant cake mix. While these products are frequently described as labor-saving devices, the perspective suggested here is that they are also (and possibly more importantly) time-saving devices. Not only does the model suggest the necessity to

view consumer durables (such as microwave ovens) as leisure-time producing devices and consumer nondurables (such as fast foods) as leisure-time producing goods, but it also focuses attention upon the entire service industry as part of the second market for time.

In addition to the obvious research potentials of establishing the time prices for goods and services that consumers will be willing to pay in the second market for time, an additional perspective is suggested. The existence of a second market for time means that an individual can add assets without giving up income. The individual still has his income but the question is one of how it is spent. If the consumer acts through the traditional market for time (first market for time), incomes are reduced because of reductions in working hours, but if the consumer purchases leisure in the second market for time, he retains more options.

### Prices in Markets for Time

Two markets for time exist for consumers for which prices exist. The price of the first market for time, which is defined to be reduction of work, is equated to foregone income. The price of time in the second market is foregone spending on nonleisure producing goods and services. Both cases involve a trade-off between money and time. In each market, money is exchanged for additional leisure time, hence time prices in each market are analytically equivalent. Which market the consumer enters is largely dependent upon the ease of market entry, that is the operational nature of the market.

The operational nature of the first market (reduction in workweek) is relatively inflexible for individual consumers. The operational nature of the second market (reduction of spending on nonleisure producing goods and services) is whether in fact consumers are aware of the actual prices of time and how much leisure time in fact will be saved. Since in each market, money is exchanged for additional leisure time, which market the consumer enters is primarily determined by the ease of market entry, i.e. the operational nature of the market, and the price per unit of leisure time in each market.

### Resource Value Inversion

The rapid spread of discretionary incomes to the masses of consumers in the industrialized economies has a profound effect upon the effectiveness of marketing strategies for time-leisure goods. The phenomenon has created what may be termed a resource value inversion which has important consequences for the development of marketing strategies.

Engulfed by the tidal wave of spending in the postwar period, business strategists have apparently neglected the fundamental aspect of consumer behavior described in the preceding pages. Namely, consumers spend from two resource budgets: a time budget and a money budget.

The fact that the importance of the time budget and the money budget may vary between consumers is called the Principle of Resource Value Inversion. Specifically, when consumers' incomes rise (as they have in recent decades), time ascends the value scale relative to money. As incomes grow, consumers place more value on their time budgets. This principle is of importance most obviously as an influence on product design and its effect on shopping patterns. It may also have considerable influence on the messages that should be used in the communications program.

Why does Resource Value Inversion (RVI) occur and how does it affect marketing strategies? One obvious reason for RVI is that on most commodities consumers must spend time as well as money. Rising incomes and more spending means that consumers must also have more time if they are to have the goods and services they prefer. RVI may also result from changing consumption patterns. Consumers may prefer proportionately more purchases that require spending larger amounts of time. The dramatic growth of the leisure industry is evidence that Americans are changing the valuation they place on resources.

Demand factors alone, however, cannot account for the Resource Value Inversion that has occurred since World War II. Business itself has contributed greatly to RVI. Since consumers traditionally have augmented time budgets by reducing working hours (First Market for Time), employers are primary suppliers of leisure. In this century workers have taken two-thirds of their productivity gains in the form of added income and one-third as added leisure.

Regaining leisure hours from employers for most workers has become increasingly difficult as the process has become institutionalized and inflexible. Consequently, workers as consumers find it harder and harder to augment leisure time as incomes have risen. Contrary to popular belief, in the postwar period consumers have not been swamped with leisure time. In fact, just the opposite has occurred: While money budgets have grown rapidly, the flow of leisure hours has slowed to a trickle.

If attention is focused upon how consumers have increased their time budgets over the past century, two trends are apparent. First, leisure gains during the sixties were small. In the decade of the sixties, reductions in working hours added only 50 hours a year to leisure time budgets compared to the century's average of 80 hours per year. It is estimated that only about 8 per cent of productivity gains during the sixties went to augment time budgets; the remaining 92 per cent went to money budgets (Moore & Hedges, 1971).

Second, the form in which leisure gains are taken has changed significantly. Whereas most of the leisure gains in the past century have been in the form of a shorter workweek (85 per cent of them), since 1950 the average workweek has remained relatively constant (Moore & Hedges, 1971). (The small declines in "average" workweek that have sometimes been reported reflect an increasing proportion of part-time workers and a decreasing proportion of farm-workers. Scheduled hours for full-time workers showed little change, averaging 45 hours in 1970.). Rather than reductions in the workweek, adding leisure time in "blocks" (paid vacations and holidays) is increasing much more rapidly than in past decades. Leisure gains since World War II can perhaps best be characterized by the statement: "Consumers Play with Blocks" -- blocks of leisure time, that is.

There is little doubt that Resource Value Inversion is a byproduct of our modern affluent economy, but what precisely does it mean for marketing strategies? By analyzing both the First and the Second market for time, several conclusions emerge. Although many other effects are possible, some of the implications include the following:

1. A growing importance for goods and services that economize on the consumers' use of time.
2. A growing importance of goods and services that require spending leisure time in "blocks".

3. A decline in the effectiveness of monetary incentives relative to leisure time incentives.

The above conclusions apply very broadly to consumer goods and services and to labor productivity as well. Referring specifically to leisure activities, there is some evidence that spending from the time budget is of more importance to consumers than spending from the money budget (Hawes, 1974; Voss, 1974). The increasing demand for blocked time could well be met by various forms of the flexible workweek (4-day week, gliding time, etc.), by increased amounts of vacation time and 3-day weekend holidays, by sabbaticals and earlier retirements (Voss, 1971). If these blocked time options are taken by American consumer-workers, the implications for marketing will be as significant or more so than for the production function.

#### Research Needed on the Markets for Time

A great deal of additional research is needed which would provide an in-depth investigation of concepts presented in this paper. Some of the more promising research needs are described below.

First, there needs to be quantitative identification of the markets for time with careful definition of the dimensions of time markets. Before questions of pricing policies, advertising effectiveness, market efficiency, and so forth can be analyzed from a time perspective, the market for time must be identified and its conceptual framework operationalized to a much greater extent.

Each division of the market for time has important implications for consumer research and for the business community as a whole. The market for time also has important implications for government agencies, however. As the dimensioning of this market is accomplished, it may become apparent that taxing policies may be influenced (since increases in income are taxed but increases in discretionary time are not).

Second, there needs to be a careful analysis of the relative attractiveness to consumers of the two time markets. Consumer surveys and experiments should investigate in which market consumers prefer to buy leisure time. An attempt should be made to assess why consumers tend to prefer one market over the other. In this context, the pricing mechanism in the two markets should be considered carefully to establish estimates of time prices.

Third, a comprehensive analysis needs to be completed defining the relationship between the second market for time and specific consumption markets. For example, the effects of time valuation should be assessed in relation to consumer durable goods, consumer nondurable goods, and service industries. While this may appear to be of most importance to the private sector, we would argue that it is of great significance to the public sector as well. Obviously, if time values are of such high magnitude that consumers are willing to purchase large quantities of disposable containers (such as in fast food industries), a significant refuse problem is created which must be solved by the government. It would be helpful to identify the probable effects upon such applications due to increasing affluence in the money budgets of consumers.

The provision of the services by the government is also an area directly affected by consumer markets for time. The government has the choice of



providing highways that will speed consumers rapidly along their way or of not doing so. Also, the government can build bridges, airports, and other facilities that will save time but it must know the relative value of such facilities to the public (Voss, 1968).

Fourth, the facilities for leisure need to be better understood as a result of an improved model of consumer time expenditures. Perhaps, the only person affluent enough (in his time budget) to attend a baseball game is a person who is relatively poor (in his money budget). Similarly, if consumers are impoverished (in the time dimension) of their daily workweek budget they may only be able to afford the time involved in tennis and may be willing to pay large amounts from their money budget for that small amount of leisure time. At the same time, if budgets of "blocked" time are increasing substantially, preferences for time-intensive leisure activities (such as golf) may be increasing dramatically by the timing of blocked time (such as on holiday weekends, retirement, and so forth). The effect on leisure-time facilities is of such importance as to demand dramatically increased consideration both in the private sector and in the public sector of the economy.

### Conclusions

The purpose of this paper is to demonstrate the need for a careful consideration of the role of time in the analysis of consumer behavior. A conceptual model is presented which emphasizes a redefinition of work-leisure model and focuses (instead) on leisure as discretionary time. Such an analysis leads to the conclusion that two markets for time exist. Not only can leisure be purchased by a reduction in the workweek (the first market for time) but leisure can also be purchased by the second market for time--goods and services which produce leisure time.

This analysis should be helpful in understanding the potential of new products and services and in the appeals that might be promoted concerning goods and services. In contrast to the first market for time, the second market for time has a number of appeals. In the first market for time, people simply work less to obtain more leisure. In the second market for time, however, people purchase more goods and services to provide more leisure. Thus, not only do they help themselves but they also contribute to increases in the Gross National Product at the same time.

### FOOTNOTES

1. Justin Voss is Assistant Professor of Economics and Environmental Science at State University of New York at Plattsburgh.
2. Roger D. Blackwell is Professor of Marketing, The Ohio State University and Vice-President, Management Horizons, Inc.

## REFERENCES

- Becker, G. A. A theory of the allocation of time. Economic Journal, 1965, 75, 493-517.
- Brennan, M.J. Theory of economic statistics. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1965.
- Engel, J., Kollat D., & Blackwell, R. Consumer behavior. New York: Holt, Rinehart, & Winston, 1973.
- Foote, N.N. The time dimension of consumer behavior. In J.W. Newman (Ed.), On knowing the consumer. New York: John Wiley and Sons, 1966.
- Hawes, D.K. An explanatory nationwide mail survey of married adult leisure-time behavior patterns and the satisfactions derived from leisure-time pursuits. Unpublished doctoral dissertation, The Ohio State University, 1974.
- Johnson, M.B. Travel time and the price of leisure. Western Economic Journal, 1966, 4, 135-145.
- Linder, S.B. The hurried leisure class. New York: Columbia University Press, 1970.
- Moore, G.H. & Hedges, J.N. Trends in labor and leisure. Monthly Labor Review, 1971, 94, 3-11.
- Pitkin, W.B. The consumer--his nature and his changing habits. New York: McGraw-Hill Book Company, Inc., 1932
- Schary, P.B. Consumption and the problem of time. Journal of Marketing, 1971, 35 (2), 50-55.
- Voss, J. The definition of leisure. Journal of Economic Issues, 1967, 1, 91-106.
- Voss, J. Evaluation of leisure time preferences in regional recreation demand. In Proceedings of Northern New York-Lake Champlain Environmental Conference, Plattsburgh, New York: 1974.
- Voss, J. Megalopolization and leisure. Natural Resources Journal, 1968, 8, 509-512.
- Voss, J. The role of leisure in economic analysis: A critique. Unpublished master's thesis, University of Missouri, 1964.
- Voss, J. Hearings on proposed adoption of a four-day, forty-hour workweek, without payment of time and one-half overtime compensation for work days exceeding eight hours. U. S. Department of Labor, 1971.



## TIME AS AN INDICATOR OF SOCIAL CHANGE AND THE QUALITY OF LIFE

John P. Robinson  
Survey Research Center  
The University of Michigan

The paper briefly summarizes research on time use that has been taking place since 1965, highlighting some aspects of this research that should be of interest to consumer research. The available national data base consists of an urban probability sample of 1244 adult respondents interviewed in the Fall of 1965 and the Spring of 1966. Several methodological studies have been conducted on smaller and more localized samples in the interim.

Major analytic efforts with the 1965-66 data have been cross-national in scope since the data collection was geared into a larger 12-nation study. Such cross-national analyses revealed far less variation than expected in patterns of daily living across societies at widely different levels of economic development. At the same time, daily life within a society behaved regularly enough to identify each society with a characteristic "life style" and these life-styles generated the provocative cross-national portrait in Figure 1.

A second direction our research has taken is toward the documentation of the changing time-use patterns in this country. Compared to earlier sociological time-use studies, our data indicated little change in the gross allocation of time devoted to work, house work and leisure. Nonetheless, Americans did seem to spend more time traveling and shopping than they had 10 or 30 years previously. However, our data clearly indicated that on the temporal dimension television must have had far more impact than any other technological advance of this century. These data also revealed Americans to be far less addicted to television than figures from the rating services, and put American mass media behavior in much clearer perspective.

More recently, our efforts have been directed toward constructing a general model of time allocation. Figure 2 illustrates a preliminary working model that has guided several multivariate analyses, such as that in Table 1 which attempts to account for variations in time devoted to house work. Table 1 indicates employment status of the woman to be the major determinant of house work, followed by presence and age of children. House work, surprisingly, varies with age but not with social class variables. None of these variables is markedly associated with male contribution to house work.

One further direction of our research has been methodological. We have compared several methods of collecting time-use data, including a field validation of the diary approach. We are also attempting to add more social psychological meaning to time use, by incorporating measures of the degree of planning, constraint, energy and satisfaction that are associated with the activities that make up daily life.

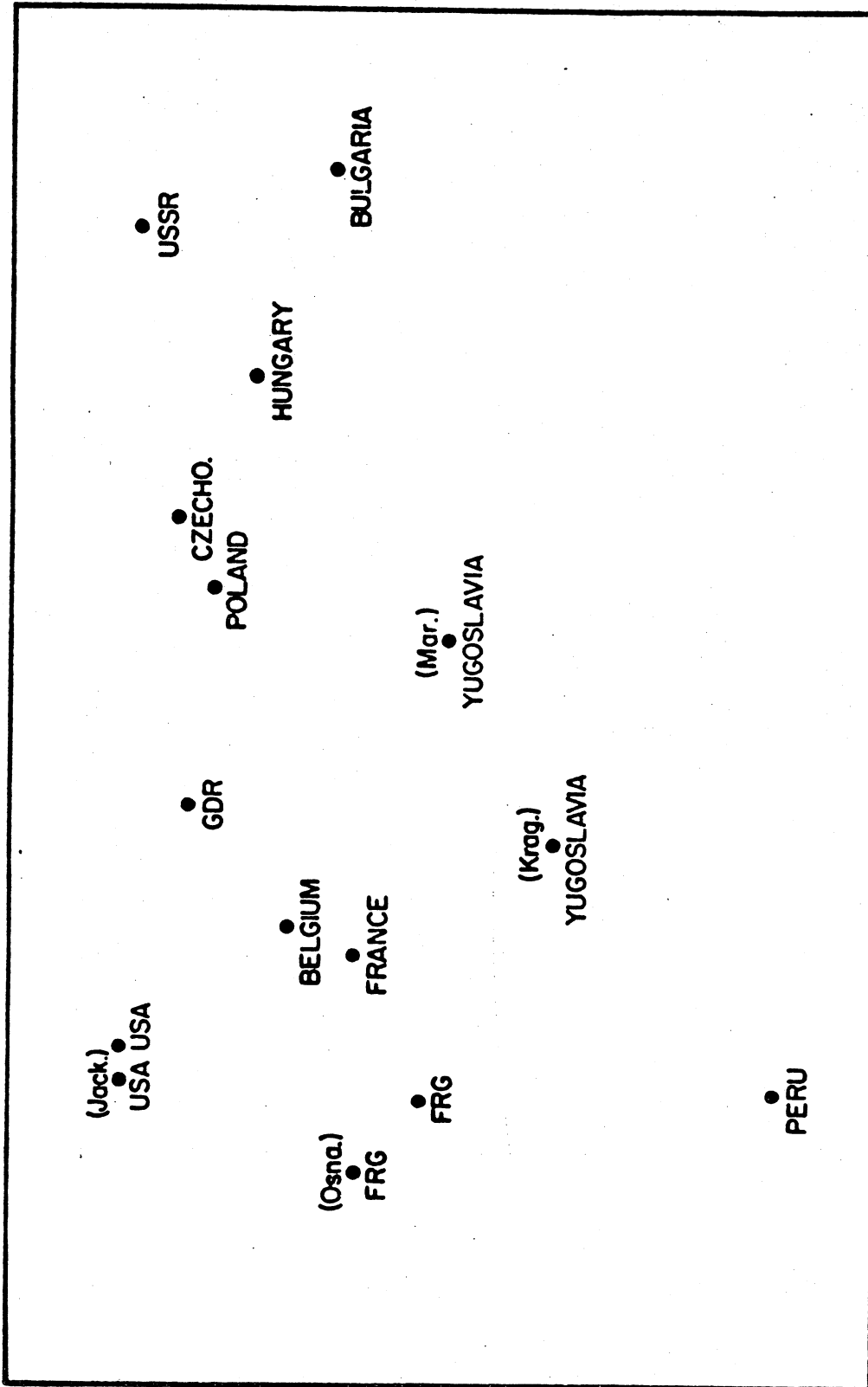
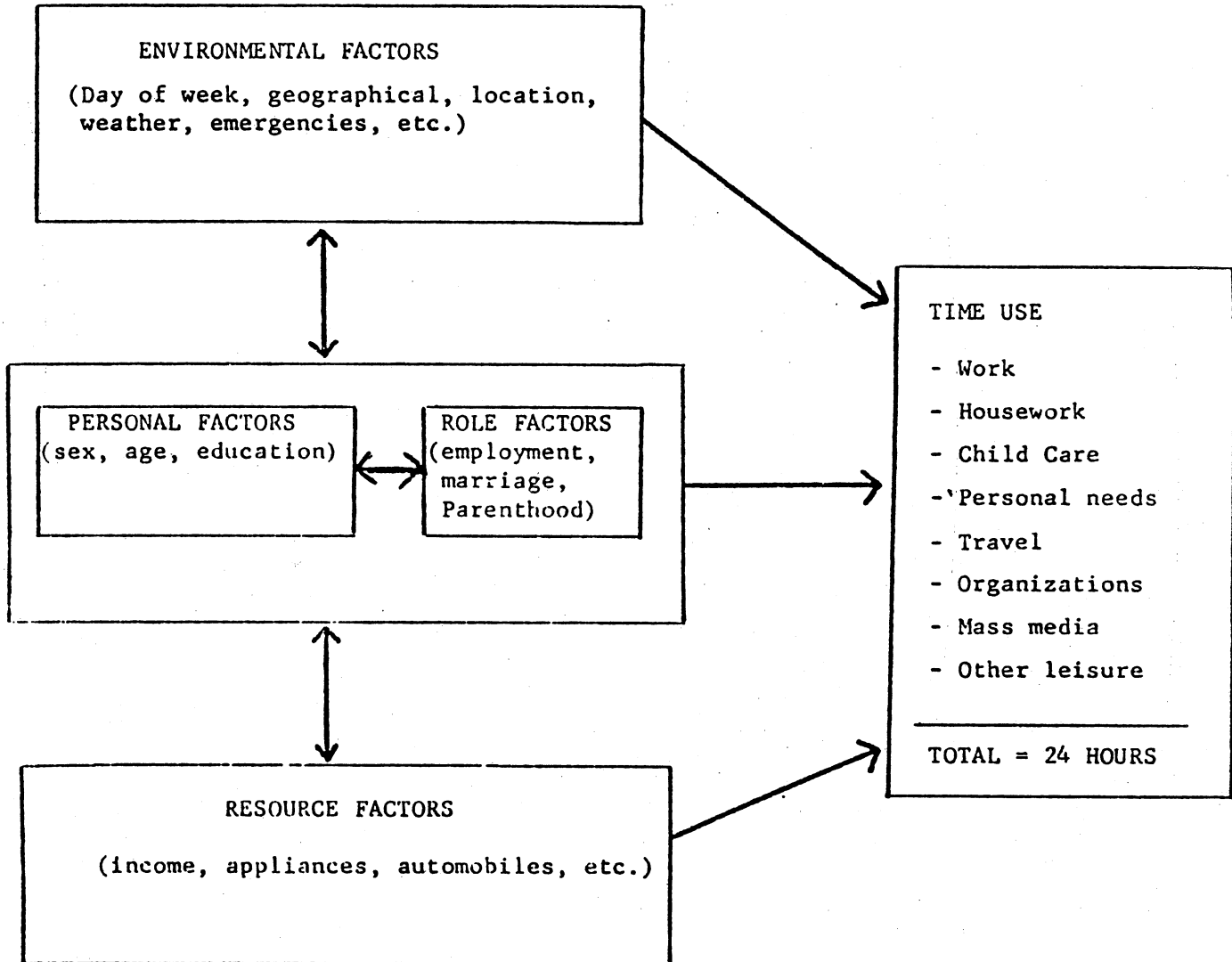


Figure 1: Two-Dimensional Solution for Site Differences in Overall Time  
Use Profiles (37 Activities)

Figure 2: Schematic Model of Factors in Time Use



	(N)	WOMEN	(N)	MEN		(N)	WOMEN	(N)	MEN	
TOTAL SAMPLE		247min		52min						
<u>PERSONAL</u>					<u>ENVIRONMENT</u>					
<u>EDUCATION</u>					<u>CITY</u>					
Grade School	(126)	261	(129)	46	Large city	(131)	272	(98)	55	
Some high school	(216)	251	(187)	53	Medium city	(156)	215	(138)	55	
High school grad	(492)	249	(304)	52	Large suburb	(185)	263	(134)	51	
Some college	(154)	223	(135)	46	Medium suburb	(186)	244	(139)	45	
College grad	(109)	253	(141)	62	Outlying areas	((42)	289	((32)	53	
<u>AGE</u>					<u>CITY SIZE</u>					
18-29	(307)	217	(241)	45	New York	(105)	231	(72)	36	
30-39	(251)	239	(198)	49	Los Angeles, Chicago	((86)	219	(67)	58	
40-49	(264)	266	(228)	55	Other large cities	(186)	256	(138)	54	
50-59	(67)	271	(233)	60	Medium cities	(182)	262	(161)	46	
<u>RACE</u>					<u>Smaller cities</u>					
White	(1013)	247	(835)	52	Jackson	(361)	248	(361)	54	
Black	(79)	240	((55)	53	<u>DAY OF WEEK</u>					
<u>RELIGION</u>					Weekday					
Catholic	(299)	242	(233)	49	Saturday	(155)	272	(113)	85	
Jewish	(42)	221	(23)	29	Sunday	(156)	216	(138)	77	
Established Prot.	(155)	251	(105)	63	<u>SEASON</u>					
Traditional Prot.	(503)	249	(418)	51	Fall	((770)	251	(611)	53	
New Prot.	((76)	255	(70)	55	Spring	(333)	237	(291)	49	
None, other	(29)	246	(53)	52	<u>RESOURCES</u>					
<u>ROLE</u>					<u>INCOME</u>					
<u>EMPLOYMENT</u>					Under \$4000					
Yes	(534)	188	(65)	50	\$4000-5999	(200)	260	(132)	55	
Student	(10)	147	(18)	151	\$6000-9999	(446)	251	(382)	54	
No	(558)	305	(19)	49	\$10,000-14,999	(215)	230	(316)	47	
<u>SPOUSE EMPLOYMENT</u>					\$15,000 and over					
Yes	(811)	255	(248)	57	(94)	260				
No	(30)	290	(509)	51	<u>PAID HELP:</u>					
No Spouse	(251)	218	(140)	51	None	(857)	248		X	
<u>FAMILY COMPOSITION</u>					Infrequent					
No Children	(422)	215	(323)	49	Part-Time	(215)	236		X	
At least					Full-Time	((19)	214		X	
one child	One	(60)	237	(51)	42	<u>WIFE'S WORK</u>				
under 4	Two	(70)	257	(71)	59	None	(561)	308	X	
years	Three	(54)	278	(45)	58	1-29 hrs	(75)	211	X	
	Four+	(74)	311	(58)	60	30-39 hrs	(98)	188	X	
All	One	(130)	250	(114)	50	40-45 hrs	(263)	188	X	
children	Two	(137)	260	(129)	54	46 hrs+	(97)	153	X	
over 4	Three	(93)	273	(68)	59					
years	Four+	(54)	290	(41)	42					

## RESEARCHING AND MODELING CONSUMER CHOICE BEHAVIOR IN URBAN TRANSPORTATION

Christopher H. Lovelock

Harvard University, Graduate School of Business Administration

Consumer behavior theory and modeling approaches can help improve managerial understanding of how consumers choose between alternative transportation modes for urban travel. This paper develops a microanalytic model of modal choice in flowchart form, clarifying the stages in the modal choice decision process for any given trip. Individual consumers are seen as trying to satisfy a particular travel need, by first specifying the characteristics of the trip itself, and then specifying the "ideal" modal attributes required for this trip. Next, the perceived characteristics of a limited number of modes are evaluated against this "ideal" solution and the consumer is posited to select that mode which provides the best match. The model explicitly recognized the impact of psychological variables on modal choice as well as the consumer's need for information.

### Public Transportation as a Marketing Problem

Recent years have brought a growing recognition of the problems raised by excessive use of automobiles as a mode of travel in urban areas. At the present time, the United States is entering what has been described by some as a "renaissance era" in public transport, spurred by concerns over pollution and congestion, urban renewal plans, opposition to new freeways and, most recently, by the energy crisis. New and improved bus and rail services are being speedily developed in numerous American cities, both large and small, assisted by major investments of public funds and accompanied by a public takeover of many privately owned operations.

The overall result has been a growing involvement by the public sector in the field of urban mass transit, with the objective of reviving dying systems and of developing new ones. But it is not enough to invest money in public transportation. The real challenge is a marketing one: how do we persuade people to use it? Transit managers simply cannot afford to lose sight of the fact that unless public transit is well patronized, many of the benefits claimed for it will never be achieved.

A brief description of the tasks facing transit planners and managers will suffice to indicate why an understanding of consumer behavior is so important in this field of activity. At the outset, it is essential that transit services be designed (or redesigned) to appeal to those who account for the bulk of car travel in urban areas, while also meeting the needs of so-called "captive" riders who either have no car or cannot drive. This task requires obtaining information on what characteristics consumers would ideally like to find in a transportation mode for various types of journeys, and then developing new or improved transit services which come as close as possible to meeting this ideal. Subsequently, the resulting services must be marketed in such a way that



consumers become aware of them, perceive them as competitive with the automobile--if they have one--and are actually persuaded to use transit. Where car owners are concerned, transit usage may require significant changes in established behavior patterns and even in lifestyles.

In order to bring about such a behavioral change, transit marketers must learn more about the criteria which people employ in making decisions relating to travel modes, as well as the relative importance of these criteria. Management will also need to know how individuals set about making decisions concerning modal choice, what information they use in making such decisions, and how they obtain it. Finally, transit managers must find the most effective means of communicating with different segments of the market for new transit services together with the most appropriate appeals to employ in seeking to change the behavior of consumers in these various segments.

### Consumer Research Traditions in Urban Transportation

Over the years, a considerable volume of consumer-oriented research has been conducted on the topic of urban transportation. However, on closer analysis, it transpires that much of this has only limited usefulness for developing transit marketing strategies. Past transportation research may be categorized into three groups:

1. Origin and Destination Studies
2. Mode Choice Models
3. Research into Modal Attributes

Origin-Destination studies are part of the standard stock in trade of the transportation planner. As the name suggests, their primary purpose is to identify the extent of travel taking place between any two or more geographic points. In their simplest form, travelers are surveyed at selected cordon points and asked where they have come from and where they are going. A note is also made of the time of the interview and the mode in which the respondent is traveling. Recently, such "O & D" studies have become a little more sophisticated, with the questions or interviewer's notations including information on the consumer's frequency of travel on specific journeys and some salient demographic characteristics. In this way, a consumer profile can be built up which may be very useful for market segmentation purposes. Repeated applications of the same survey instrument at periodic intervals can show changes in travel behavior over time, possibly reflecting changes in transportation system characteristics (e.g. increased fares, development of new routes, etc.). However, O & D studies ignore attitudinal factors completely.

From the data collected by such studies, transportation researchers have developed a substantial number of quantitative models designed to forecast "modal split"--that is, the proportion of trips accounted for by each of the modes available in a given travel corridor. A huge number of Mode Choice Models have been developed and are described in review papers such as Bock (1968) and Hartgen (1970). Typically, such models contain a strictly limited number of independent variables. Many, for instance, were designed simply to demonstrate the impact on transit's share of the travel market of changes in transit fares. These models resulted in such managerially unhelpful findings as a conclusion that the only way to increase transit's market share in a major midwestern city was to have a negative fare for each passenger of 35 cents. The other widely used variable, often employed in conjunction with price, was travel time. The best of the economic models of modal choice were probably the utility models grounded in Lancaster's (1966) general

theory of consumption. These models assumed that travel, as a derived demand, was a negatively valued utility and posited that consumers would select that mode of transportation which represented the least disutility. The disutility of travel by a given mode was defined as the sum of the disutilities associated with each modal attribute specifically included in the model. One significant problem with this utility model is that it assumes disutility to be a linear function of an attribute's magnitude, thus ignoring threshold levels of acceptability and diminishing (or increasing) marginal utility.

The third category, Research into Modal Attributes, clearly has a useful role to play in both transportation system design and model construction. The objectives of this type of research are, first, to determine how consumers rank different attributes of transportation modes and second, to discover how well specific modes are perceived as performing on these desired characteristics. These studies gather data through the use of large-scale consumer questionnaires, and typically employ Likert-scaled items or paired comparisons to arrive at rankings of attributes. Several major research studies (Paine *et al.*, 1967; McMillan and Assael, 1968, 1969; and Golob *et al.*, 1972), employing varying methodologies and sample populations, have yielded reasonably consistent findings concerning the characteristics required by consumers in a transportation mode and the relative importance which they attach to these characteristics. In order of importance, the principal characteristics which consumers seek in a mode of transportation appear to be safety, reliability, time savings, cost, convenience and comfort. However, additional findings showing consumers to rate car travel higher than transit travel on virtually every attribute are open to question, in light of the fact that substantial numbers of people in urban areas with a choice of modes do elect to use public transportation, as well as on the grounds of possible bias in questionnaire phrasing and inclusion of non-representative consumers in the sample population (Lovelock, 1972).

#### The Need for a Consumer Behavior Approach to Modal Choice Modeling

From the foregoing discussion, it becomes clear that existing transportation research and modeling techniques overlook many considerations and concepts which are widely used by marketers. For a start, modal choice models ignore the potential impact of several key elements in the marketing mix which may influence consumer behavior. Although the issues of product characteristics and price levels are addressed, the approaches employed suffer from significant shortcomings, not least in the use of a purely economic treatment of pricing strategy. Completely missing is any recognition of the role which promotion can play in influencing modal choice.

A number of transportation researchers have pointed to the limited explanatory power of models based upon economic analysis and have emphasized the need to develop a better understanding of the ways in which consumers arrive at modal choice decisions (Sommers, 1970; Le Boulanger, 1971; Horton, 1972). In particular, growing interest is being shown in learning more about how behavioral variables such as attitudes relate to modal choice (Hartgen and Tanner, 1971; Allen and Isserman, 1972).

Yet another shortcoming of previous modeling approaches is that most modal choice research has tended to be macroanalytic in nature (i.e. focusing on the behavior of large groups of travelers) and relatively little is known about the individual consumer's decision process or how his decisions might be influenced by transportation managers. In the balance of this paper I propose to outline a marketing-oriented approach to modeling modal choice, derived from research into brand choice for consumer convenience products.

## Microanalytic Modeling of Consumer Behavior

Marketing researchers have, of course, devoted considerable attention to development of models of consumer behavior. The greatest emphasis in modeling consumer choice has been given to convenience products, notably branded food items. I believe that certain of the approaches employed in modeling the brand choice decision process can usefully be adapted to decisions involving frequent selection of other types of consumer products or services, including urban transportation alternatives. A particular advantage of consumer behavior modeling, from the perspective of transportation research, is that such models often explicitly incorporate social-psychological variables, enabling one to examine the role of attitudes, perceptions, etc., in the modal choice decision process.

As noted earlier, most economic models of modal choice have been macro-analytic in nature. This focus on the behavior of large groups of consumers severely detracts from the model's ability to explain behavior and thereby provide insights into ways of changing existing behavior patterns. The focus here, therefore, will be on a microanalytic approach, emphasizing the decision process of individual consumers in selecting a mode for a particular trip.

Microanalytic simulation is a relatively new approach to modeling consumer behavior, embodying both economic and behavioral concepts. As noted by Eskin and Venkatesan (1974), these models describe the behavior of individual decision makers rather than group or market level decision processes. They also involve a specification of causal structure which seeks to explain why various decisions are made, rather than just claiming ability to forecast outcomes given knowledge of environmental conditions. A key characteristic of such models is that they simulate the actual decision-making process. The procedure used for generating forecasts and searching for optima is to run the model, as opposed to the more conventional process of solving models by analytical methods. This simulation approach is generally necessitated by the complex and interactive nature of the model's structure.

Consumer decisions are defined in terms of observable behavior, such as choice of a product or brand. Typically, these decisions are determined by some form of attitude construct, based upon consumer preference for a particular product. This preference is derived from the consumer's perception of the benefits to be derived from consumption or use of the product in a specific situation, and presumably reflects the match between his/her specific needs and the perceived attributes of each of the products or brands under evaluation.

The process of model development can be broken into two stages: description of the consumer's decision process and then quantification of the inputs. In its descriptive format, the model can be expressed as a flow chart of the decision process; this shows the various steps involved in making a decision, identifies the alternatives available and specifies the interactive nature of the model's structure. When quantified and operationalized, microanalytic simulation models typically exist in the form of a computer code, capable of producing numerical outputs such as forecasts of the number of persons purchasing specified brands under given conditions. Usually, these models are stochastic in nature, since uncertainty components play an important role in constructing the model. The characteristics and behavior of consumers, as well as the environmental conditions facing them, are generated within the model through probabilistic processes, based upon data derived from market research and managerial judgement.

Eskin and Venkatesan went on to demonstrate how a simple microanalytic simulation was successfully developed to illustrate the consumer decision flow involved in selecting a particular brand of cake mix from among competing alternatives, and subsequently operationalized to enable management to evaluate the impact of alternative market events and strategies.

This cake-mix model is described in greater detail in Eskin and Lovelock (1971, 1972). In this model, consumers are seen as first defining the occasion at which the resulting cake will be served, thus determining some of the attributes which they will seek in the product. The model limits these attributes to four key dimensions, two of which--price and quality--are treated as interrelated. Each consumer has her own "ideal" product in mind, based on the serving occasion for which it is intended and her personal preferences. Likewise, she also forms a perception of the characteristics of each available alternative, which is open to modification through promotional activity. Alternative brands are then matched against the "ideal" and it is hypothesized that she will purchase that brand which best meets or exceeds her requirements. The model provides for ties between brands and also posits that if no brand matches her ideal, the consumer will either change her requirements or else not purchase at all.

#### Modeling the Individual's Modal Choice Decision Process

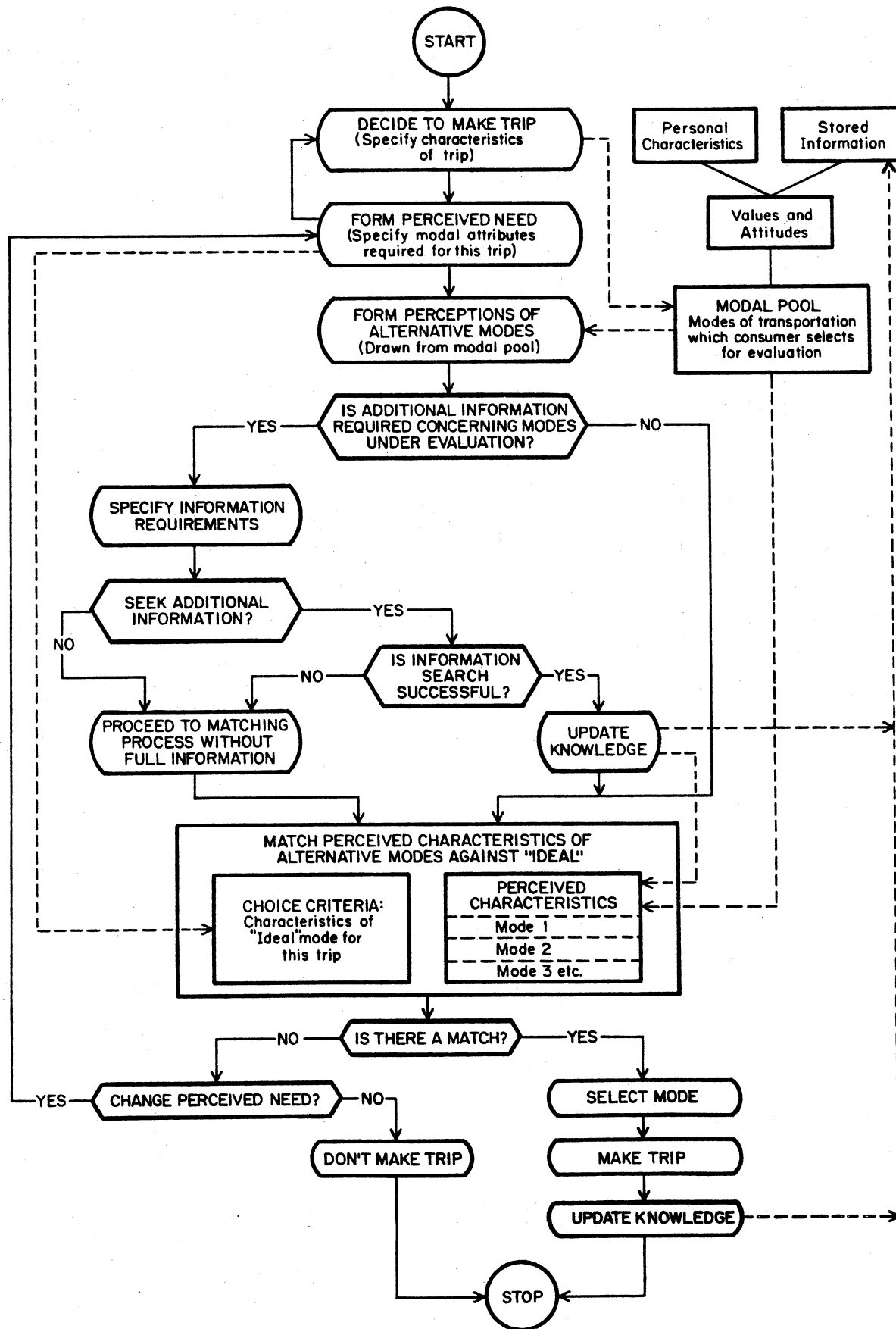
The same basic concepts underlying the cake mix model are used here to develop a behavioral model of modal choice which, I believe, provides fresh insights to the problem of marketing urban transit. The model has its antecedents in the consumer behavior simulation models developed by Amstutz (1967) and Claycamp and Amstutz (1968), as well as drawing on work by Engel, Kollat and Blackwell (1968).

Translating a model designed to study brand choice into one designed to study choice of transportation modes is simpler than might first appear. The "product classes" in this instance are represented by trips for different purposes (e.g. work and non-work trips) for which, as noted by Paine *et al.* (1967), consumers tend to have differing choice criteria. The "brands," meantime, are represented in this scheme of things by the alternative public and private modes available to the consumer. Choice of a transportation mode shares a very significant characteristic with the selection of branded food products, namely the frequency with which the choice has, theoretically, to be made. This does not mean, of course, that a different decision will be reached each time a journey is made, nor even that the consumer will necessarily think very hard about the decision, which may well have become routinized. But then, consumer loyalty and habitual purchase of an established brand can be a problem, too, for many convenience goods marketers seeking to enlarge their market share in the face of entrenched competition.

The focus here is on the first stage of model development, namely a description in flow chart form of the decision process through which the consumer is posited to go in choosing a mode of transportation for a particular journey. However, I shall also discuss some of the issues related to quantification of inputs. In this model, displayed in Exhibit 1, individual consumers will be seen as trying to satisfy a particular travel need by first evaluating alternative modes of transportation, then matching the perceived characteristics of available modes against an "ideal" solution, and finally selecting that mode which provides the best match (if such exists).

Exhibit 1

THE MODAL CHOICE DECISION PROCESS FOR A SPECIFIC TRIP



### Decide To Make Trip

Transportation is, of course, a derived demand, and so the modal choice decision in the model is preceded by the consumer's decision to make a journey which will satisfy other needs. Each trip has specific characteristics of its own which will have a major impact on the consumer's choice of a mode of transportation. Consequently, the model begins by calling for the consumer to specify the characteristics of the trip, notably the origin, the destination and the purpose of the trip; other important considerations might include the number and ages of others (if any) in his party, the time at which the trip was to be made, the prevailing weather conditions, etc. In its simplest form, this component of the model might consist of three categories, not dissimilar to the "Select Use Category" in the cake mix model:

- \*Work trip
- \*Shopping trip
- \*Social/Recreational trip

### Form Perceived Need

Having specified these trip characteristics, the consumer next moves to formulate his perceived needs and specify the modal attributes required for this trip. They constitute an "ideal" solution to his transportation needs--the choice criteria against which he or she will subsequently evaluate the various alternatives which may be available. On the basis of previous transportation research, the key attributes to include would seem to be:

- \*Point-to-point travel time
- \*Price
- \*Safety
- \*Reliability
- \*Convenience
- \*Comfort

Additional market research will be needed to develop appropriate indices of what might be either grouped quantitative data (e.g. price in cents per person, travel time in minutes) or verbalized ratings of the other characteristics. One advantage of using grouped data in this way is that it gets around the problem of purely linear functions: thus, price could be expressed as 10-14¢, 15-24¢, 25-29¢, 30-40¢, 50-74¢, etc., reflecting varying demand elasticities at different price levels.

Although a review of transportation research findings suggests that consumers are willing to trade off modal attributes against one another (e.g. they will pay a premium fare for a faster or more comfortable service), my own research suggests that there are certain threshold levels of acceptability on specific attributes which may serve to disqualify the entire mode (Lovelock, 1972). Consequently, it is necessary to include in the coding not merely the consumer's "ideal" value for each attribute but also his "threshold" value. This might consist of a maximum acceptable price or travel time, a minimum level of comfort and safety, etc.

### Form Perceptions of Alternative Modes

Now the consumer is ready to begin the search and evaluation process which will lead to a final decision on a choice of mode. The characteristics and performance of each mode can be described in the same terms as perceived need. Thus it can be given a rating in terms of travel time on a particular

journey, a price level, a comfort rating and so forth. What is important is that these are not objective ratings as determined by an outside expert, but the subjective perceptions of an individual consumer. As such, they may be far removed from reality, particularly if the consumer is poorly informed or has little experience of using the mode in question.

It is entirely possible that certain modes may be entirely absent from the evaluation process. The model makes explicit provision for this through what is termed the MODAL POOL. In this pool will be found all the alternatives which are perceived as offering a potential solution to the transportation problem raised by the trip in question (note the dotted line feeding information on the type of trip to the MODAL POOL). Each trip decision requires a new search of the pool, since the perceived alternatives are themselves a function of the trip characteristics. However, the pool does not necessarily include all feasible alternatives--only those which the consumer selects for evaluation. The selection process is a function of two factors, PERSONAL CHARACTERISTICS and STORED INFORMATION, as modified by VALUES AND ATTITUDES.

PERSONAL CHARACTERISTICS refer to demographic, physiological and personality factors: for instance, car ownership, stage in life cycle, state of health and underlying motivations may all have an important impact on the modes which people evaluate for a particular trip. STORED INFORMATION, representing past experience and awareness of the availability of alternative modes and routes, also constitutes an important determinant of the modes which are to be brought forward for further evaluation. Acting as a filter to the modal pool are VALUES AND ATTITUDES. Individuals may be perfectly well aware that there is a bus service available, but their generalized attitudes towards buses may be such that they will not even consider this mode as an alternative.

At this point, then, the consumer brings forward one or more selected modes from the MODAL POOL for further evaluation and forms perceptions of the characteristics of each.

#### Additional Information

The immediate question now is whether the consumer has sufficient information to make a decision. It may be that (s)he is thinking about taking a train; but doesn't know where the station is, what the schedules are or how much it costs. Perhaps (s)he is also considering driving and is unsure how long to allow for the trip or whether parking is available at the other end. If the consumer already has enough information and cannot be bothered to search for more (s)he can proceed straight to a decision. Alternatively, the consumer is positioned to specify the additional information required.

This information search may or may not prove successful. Unfortunately, an unsuccessful outcome to the information search is particularly likely to result in the case of public transportation, since many transit companies have very poor information services or are understaffed, so that a phone call is often met by a continuous busy signal. If consumers elect not to search for additional information even when they need it, or if they search unsuccessfully, then they will proceed towards making a final decision without the benefit of full information--with the accompanying risk that it may not be the best decision. However, obtaining the needed information permits consumers to update their knowledge and also record what has been learned under STORED INFORMATION for future reference.

### The Matching Process

The consumer is now ready to compare the perceived characteristics of each of the alternative modes against the previously defined choice criteria for this trip. It is hypothesized that he or she will select that mode whose perceived attributes (i.e. cost, convenience, comfort, etc.) meet all threshold requirements and perform best in matching (or exceeding) the attributes of the ideal, but allowing for trade-offs between different characteristics. However, there is no guarantee that a satisfactory match will occur, even when only one mode is under consideration for the trip. If a match results, one would expect the consumer to select a mode, make the trip and then update his knowledge on the basis of the experience. However, if no match results, two alternatives remain open--either not to make the trip at all or to change these perceived needs in the hope of subsequently achieving a match.

A change in needs can have one of three outcomes. The first is represented by a decision to modify the ideal, so that the consumer is prepared (say) to accept a slower or more expensive trip, use a less comfortable vehicle, etc. Second, would-be travellers can modify the nature of the trip itself, so that it is effectively no longer the same trip--they can leave later, visit a different destination, not take children along with them, decide to forego one of the purposes of the trip, etc. The model provides a feedback loop between FORM PERCEIVED NEED and DECIDE TO MAKE TRIP which, it will be noted, re-establishes contact with the MODAL POOL. This action makes possible a third solution, namely another search of the pool and introduction of new modes into the modal decision process.

### Discussion of the Model

The value of this model is twofold. The flow chart alone serves to highlight the information we need to know about individuals and their behavior if we are to be able to understand and influence their modal choice decisions. It is hoped that further research will make it possible to quantify the model and use it also as a predictive tool.

A review of the flow chart suggests that we need, first, to know something about each individual, notably demographic characteristics (including, of course, vehicle ownership), physical condition, and personality. This information serves both as a guide to the type and number of trips (s)he is likely to make on any given day and also acts to define the constraints surrounding the modal choice decision process (e.g. a 14-year old cannot legally drive; a person in a non-car owning household is unlikely to have a car available as a possible alternative; a disabled person may not be able to use public transportation, etc.).

We also need to know what information is stored away in the consumer's mind concerning transportation alternatives--which is a function of past experience, exposure to advertising, word of mouth, etc. An understanding of this attitude and value structure can help us understand why one individual's decision process may be sharply different from that of another with broadly similar demographic characteristics. The concept of the MODAL POOL proposed in this model suggests that modal choice is a two-stage process, with the first stage representing a possibly subconscious filtering process by which only a restricted number of possible alternatives are advanced for conscious evaluation. In many instances, it is believed that only one mode, the private automobile, emerges from the pool, with the result that for repetitive journeys the entire modal choice process may be undertaken almost subconsciously.



The consumer's evaluation process focuses our attention on four considerations. The first is the development of an "ideal" modal solution to the trip in question, based upon the characteristics of that trip; this recognizes that a journey between the same origin and destination may not always constitute the same "trip" in the consumer's mind and helps to account for the fact that the same person may not always use the same mode each day for the same journey (Lovelock, 1972). The second point of interest is the need the consumer may have for additional information in order to be able to make a thorough evaluation. If a consumer recognizes such a need but considers that obtaining the information would represent too much of a "hassle," then the choice will be made on the basis of imperfect information and the decision may go, by default, in favor of the mode about which (s)he is already best informed. This has important implications for transit marketers, in that a car owner is likely to be much better informed about driving alternatives than about transit alternatives.

The third element of interest is the matching process itself, where the perceived characteristics of the competing modes are compared against those of the "ideal." Most existing modal choice research has emphasized modal characteristics as the key determinants in the modal choice decision process, but has used a very limited number of attributes. Such research has also tended to employ aggregative models which treat the characteristics of alternative modes as "facts," rather than recognizing differences in perceptions, and also to assume that the choice criteria (i.e. the ideal) remain constant for broadly defined types of trips. Inevitably, such models tend to regard the characteristics of alternative modes as fixed but largely unrelated physical attributes which can only be altered by making physical changes (e.g., introducing new vehicles, speeding up the service, reducing fares, etc.). By contrast, behavioral theory would argue that each individual can have different needs and perceptions and that it may be possible to change perceptions of key characteristics through persuasive communications. Alternatively, a physical change in one attribute (say putting new seats in a vehicle to improve comfort) might serve to influence consumer perceptions of other attributes through a "halo effect."

The fourth point of interest in the model is the updating of information which takes place at two different stages--following a successful information search and again following actual usage of a mode for a particular trip. This focuses attention firmly on the impact which awareness of alternatives and past experience can have on the modal decision process. Another input to updating knowledge (not shown in the model) would be represented by receipt of messages concerning transportation alternatives and their characteristics. In the case of public transportation, these might be expected to result from advertising campaigns or public information programs. Promotional campaigns might also attempt to change attitudes in order that transit would be included in the MODAL POOL as an alternative at least worthy of evaluation.

#### Future Development of the Model

At this point, the next stage is to attempt to quantify the model. Initially, it will be advisable to develop a very simple model, tailored to a specific modal choice situation in a given location. Microanalytic models can quickly become immensely complex and expensive to run in terms of computer usage. One of the virtues of the cake-mix model described earlier is its simplicity--and yet it does seem to produce meaningful results. In quantifying and operationalizing the model, the following tasks will be necessary:

1. Select a few basic types of trip for a specific route
2. Select key modal attributes and develop appropriate indices for each
3. Identify key segmenting variables and the proportion of consumers in each segment
4. Specify "ideal" and "threshold" values for each attribute by each consumer segment and for each trip type
5. Specify different perceptions of each attribute by each consumer segment for each trip type and for each mode of transportation included in the model

Market research will be necessary to collect new data specifically tailored to the requirements of the model. With the aid of such data, a "real-life" situation can be simulated and sensitivity analysis conducted on the inputs to see what impact changes in either the ideal or the perceived attributes would have on modal choice behavior. Such changes could then be replicated under real-world conditions, through physical changes in transit attributes in a restricted location or promotional campaigns by the transit authority designed to alter perceptions regarded as non-representative of reality. Subsequently, the model could be validated by comparing the modal split predicted by the output with the known behavior of the consumer population being simulated. By varying the inputs, it should be possible to fine-tune the model and thus improve its accuracy.

#### REFERENCES

- Allen, W., and Isserman, A. Behavioral modal split. High Speed Ground Transportation Journal, Summer 1972, 179-199.
- Amstutz, A. E. Computer simulation of competitive market response. Cambridge, Mass.: The M.I.T. Press, 1967.
- Bock, F. C. Factors influencing modal trip assignment. Washington, D.C.: National Cooperative Highway Research Program, Report No. 57, 1968.
- Claycamp, H. J., and Amstutz, A. E. Simulation techniques in the analysis of marketing strategy. In F. Bass, C. King, and E. Pessemier (Eds.), Application of the sciences in marketing management. New York: John Wiley & Sons, 1968.
- Engel, J. F., Kollat, D. T., and Blackwell, R. D. Consumer behavior. New York: Holt, Rinehart & Winston, 1968.
- Eskin, G. J., and Lovelock, C. H. Hinesbury Mills (A), (B). In G. J. Eskin and D. B. Montgomery (Eds.), Cases in computer and model assisted marketing. Vol. 2. Cupertino, Calif.: Hewlett-Packard, Inc., 1974.
- Eskin, G. J., and Venkatesan, M. A micro analytic simulation of consumer behavior: Economic and behavior theory antecedents. In S. Ward and P. Wright (Eds.), Advances in consumer research. Vol. 1. Urbana, Ill.: Association for Consumer Research, 1974.
- Ewing, R. H. Psychological theory applied to mode choice prediction. Transportation, 1973, 2, 391-410.

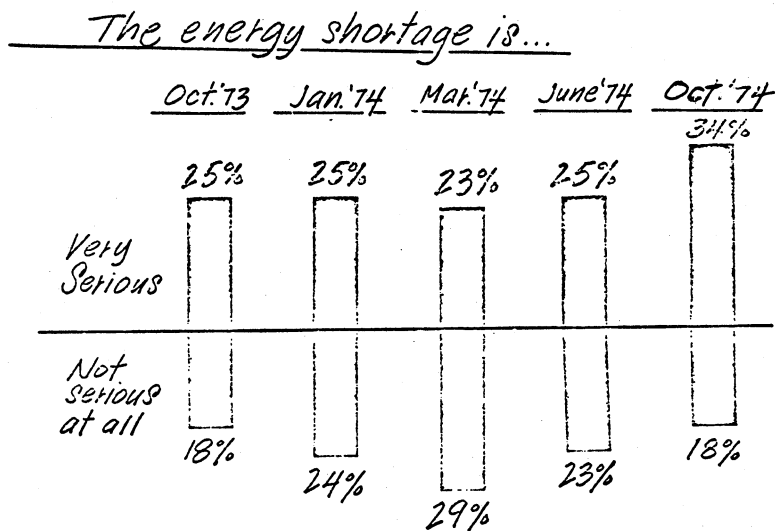
- Golob, T.F., Canty, E.T., and Gustafson, R.L. An analysis of consumer preferences for a public transportation system. Transportation Research, 1972, 6, 81-102.
- Hartgen, D.T., and Tanner, G.H. Investigations of the effect of traveler attitudes in a model of modal choice behavior. Highway Research Record, 1971, 369, 1-14.
- Hartgen, D.T. Mode choice and attitudes: A literature review. Albany, N.Y.: New York State Department of Transportation, 1970.
- Horton, F.E. Behavior models in transportation planning. Transportation Engineering Journal, Proceedings of the American Society of Civil Engineers, May, 1972, 411-420.
- LeBoulangier, H. Research into the urban traveller's behaviour. Transportation Research, 1971, 5, 113-125.
- Lancaster, K.J. A new approach to consumer theory. Journal of Political Economy, 1966, 6, 132-157
- Lovelock, C.H. Consumer oriented approaches to marketing urban transit. Doctoral dissertation, Stanford University, 1972. Published by National Technical Information Service, Springfield, Va. Report No. PB-220 781/9.
- McMillan, R.K., and Assael, H. National survey of transportation attitudes and behavior. Phases I and II. Washington, D.C.: National Cooperative Highway Research Program Reports No. 49 (1968) and No. 82 (1969).
- Paine, F.T., et. al. Consumer conceived attributes of transportation: An attitude study. Springfield, Va.: National Technical Information Service, No. PB-176 485, 1967.
- Ray, M.L., and Leavitt, C. Introduction: Workshop on microanalytic simulation and behavior theory. In S. Ward and P. Wright (Eds.), Advances in Consumer Research. Vol. 1. Urbana, Ill.: Association for Consumer Research, 1974.

THE EFFECTS OF ENERGY AVAILABILITY AND COSTS  
ON CONSUMER ATTITUDES AND BEHAVIOR

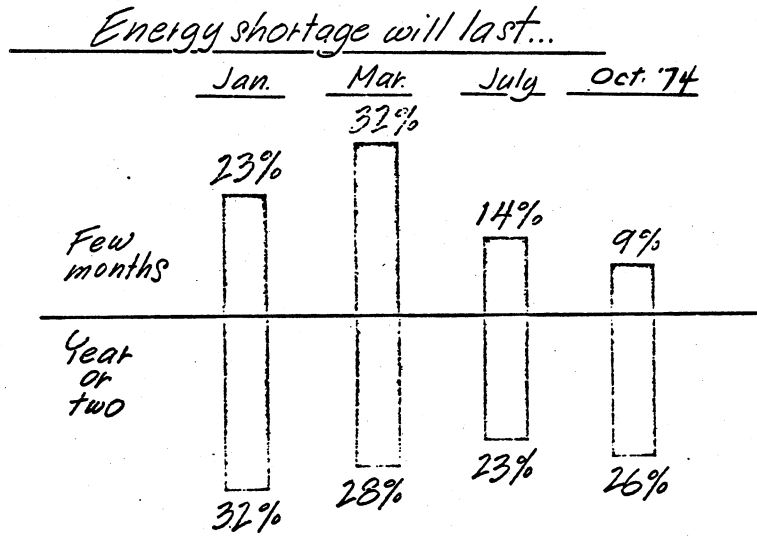
Harry W. O'Neill  
Opinion Research Corporation

Since January of this year, Opinion Research Corporation has been conducting a research program to identify and monitor changes in attitude and behavior resulting from the energy shortage and rising energy costs. The research is nationwide in scope and conducted by way of telephone interviews. Initially, 400 interviews were conducted for each wave of the program -- a wave being a two-week period. Since the middle of August, the number of interviews has been increased to 600 per wave. This paper presents some of the findings from this research program.

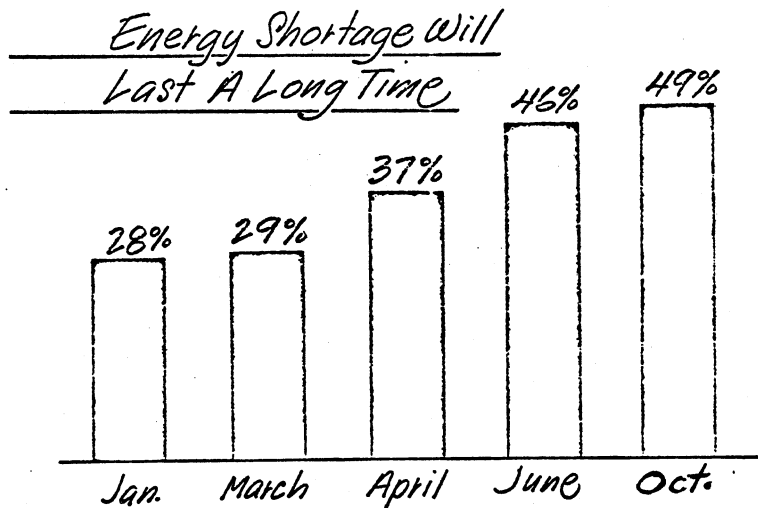
Looking at the energy shortage overall, notice that the latest findings at the right of the chart show that the percentage of the public who believe the energy shortage is very serious is about twice the percentage who believe it is not serious at all -- 34% vs. 18%. Back in June, July, and August of this year, the two groups were about equal -- about a fourth of the public believing the energy shortage is very serious and a fourth believing it's not too serious. When the lines at gas stations started to disappear -- back around March -- there was a considerably higher proportion of the public who felt the energy shortage wasn't serious. But as people have seen gas prices climb -- even though there has been no problem with availability -- there has been a decided drop-off in the number of people saying the shortage is not serious.



We have been asking people whether they believe the energy shortage will last only a few months, a year or two, or a long time. As you can see, only about one person in ten believes the energy shortage will be over within a few months, and another fourth hold that the shortage will last a year or two.



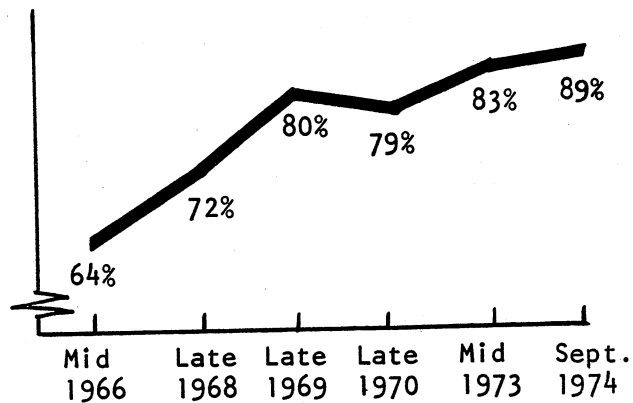
On balance, people believe that the energy shortage will be with us for a long time. And the proportion of the public who hold this point of view has been steadily increasing since January. Today, about half the public believes that the energy shortage is not a short-term phenomenon.



It is virtually impossible to separate attitudes and behavior related to the energy shortage from those related to inflation. Both of these problems are having a profound impact on our whole way of life. In a sense, double-digit inflation is moving us from an economy of more and more to an economy of less and less -- from an impulse buying economy to an economy of selective purchasing. The importance of these emerging trends resides in the fact that they may well be destined to have a significant impact on the kinds of products and services marketers can expect to sell in the future.

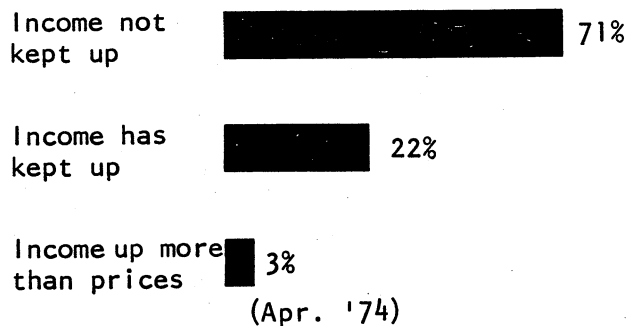
The percentage of people who say they have been hurt by inflation has reached an all-time high. Almost nine persons in ten report they are feeling the bite of inflation -- up sharply from 1966.

### Hurt By Inflation



Furthermore, seven people out of ten report that their income has not kept up with rising prices. Only a fourth are in the fortunate position of staying even or a bit ahead.

### Income Vs. Rising Prices



And it isn't just the people in the low-income brackets. Here are the same data shown by level of income. You can see that even those in the higher income brackets say that they are falling behind. Obviously, there is a relationship

with income -- but 54% of the people with family incomes of \$15,000 or over say that their income has not kept up, while only about a third say their income has kept up.

Income vs. Rising Prices  
by Family Income

	<u>Income Has Kept Up</u>	<u>Income Has Not Kept Up</u>
Under \$5,000	9%	83
\$5,000-\$6,999	22%	71
\$7,000-\$9,999	17%	79
\$10,000-\$14,999	28%	66
\$15,000 or over	35%	54

(Apr. '74)

When people are asked to look ahead to next year and predict what will happen to their family income relative to prices, over four persons in ten believe their income will fall behind while half expect to stay about even or perhaps have their income increase a bit relative to prices.

Family income  
relative to prices

Increase more	12%
Stay about even	38
Fall behind	43
No opinion	7

(Oct. '74)

All of this, of course, is reflected in how people shop. We have been monitoring since 1968 whether rising prices affect shopping habits. You can see continuous movement upward to the present time, where seven Americans in ten report that rising prices, indeed, have caused them to change their shopping habits.

When people are asked what specific changes have occurred in their shopping habits because of inflation, five behaviors are mentioned by about half or more of the public:

There is an increase in shopping at more than one store in an effort to obtain the best price.

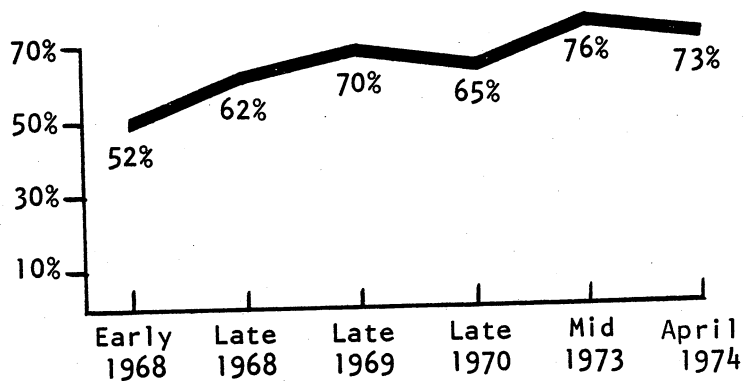
There is more buying of store brands rather than national brands.

There is an increase in the purchase of larger economy-size packages.

There is an increase in the shopping at discount stores.

There is more payment in cash rather than by charge account.

### Rising Prices Changed Shopping Habits



All in all, higher prices are having an important impact on the perceived value of key goods and services -- a few of which are shown here. For a whole range of products and services, people were asked which they believe are providing them with the least value for the money spent. Note that in the case of gasoline, there is a 26 percentage point difference between the current results and those for 1968. Of all the products and services evaluated, gasoline showed the greatest change in terms of least value.

### Least Value for the Money

	Late 1968	Apr. 1974	Points Diff.
Grocery products	30%	42%	+12
Gasoline	15	41	+26
Cosmetics, toiletries	31	23	- 8
Autos	19	21	+ 2
Prescription drugs	29	16	- 7
Electricity	5	15	+10



Most people report having been personally affected to some degree by the energy shortage; and as the price of various forms of energy has increased, the number of people reporting being personally affected a great deal has been trending up.

Personally affected by  
energy shortage ...

	<u>March</u>	<u>July</u>	<u>Sept.</u>
Great deal	15%	20%	22%
Some	34	31	35
Only a little	34	32	26
Not at all	17	16	17

Very few people report experiencing any difficulty in buying gasoline. There is no problem with availability.

Difficulty in Buying Gasoline

	<u>March</u>	<u>Oct.</u>
Great deal	9%	0%
Some	19	3
Very little	22	5
None	49	90

However, cost is now an important factor. Back in March, about two-thirds of the public reported that price was not a factor in their purchase of gasoline. Today, about half the public say it is a factor and they are looking around for the lowest price available.

Looked Around For Lowest Price  
Last Time Bought Gas

	<u>March</u>	<u>July</u>	<u>Oct.</u>
Looked around	33%	51%	49%
Price not a factor	65	48	47

People find it hard to accept what has been a steady increase in gasoline prices; and six persons in ten are of the opinion that, compared to most other things they have to buy, they find the price of gasoline to be unreasonable. As a result, the public is rather unsympathetic to the supposed problems of the oil industry.

Price of Gasoline  
Compared To  
Most Other Things

Unreasonable	60%
Reasonable	38
No opinion	2

(Oct. '74)

Not only do people feel that the price of gasoline is unreasonable, but very few expect any decrease in price in the near future. As a matter of fact, more people believe the price will go up than believe it will remain at its current level.

Over next three months,  
price of gasoline will ...

Increase	43%
Not change	38
Decrease	10
Don't know	9

(Oct. '74)

Six persons in ten report that they are now using their car less because of the price of gasoline.

Using car less because  
of price of gasoline ...

Yes	60%
No	37
Don't know	3

(Oct. '74)

And two-thirds say that if the price were to reach 70¢ a gallon there would be even a further decrease in their driving.

At 70¢ per gallon ...

	<u>June</u>	<u>Oct.</u>
Would decrease my driving	75%	69%
Would not	21	25
Not sure	4	6

The energy situation has always been rather confusing to the public. People have never been sure where to put the blame. Oil companies, however, consistently have been named more often than any other group or institution as being most responsible for the shortage.

#### Most Responsible for Energy Shortage

	<u>Jan.</u>	<u>Sept.</u>
Oil companies/oil industry	29%	26%
Federal government	19	11
President Nixon	14	8
Everyone	10	12
Public/consumers	7	11
Big business/business leaders	4	11
All others	15	23
Don't know/no opinion	15	19

In general, the public is dissatisfied with public, business, and governmental efforts to relieve the energy shortage. Furthermore, public dissatisfaction is increasing. And, on this issue, the honeymoon period for President Ford is over. In our latest measurement, nearly half the public say they are not satisfied with his efforts regarding the energy situation. A month before, only 25% of the public expressed dissatisfaction.

#### Not satisfied with steps to relieve energy shortage taken by ...

	<u>April</u>	<u>Oct.</u>
Oil companies	57%	63%
Nixon/Ford	51%	46%
Congress	41%	56%
Auto companies	37%	48%
Public	19%	36%

Since the advent of the energy shortage as a major public issue, oil companies have been devoting a considerable amount of effort to advertising that discusses the energy shortage, its causes, and what they as companies are doing about it. About six persons in ten report being aware of such oil company advertising.

Aware of oil company advertising about energy shortage and what they are doing about it	58%
Not aware	42
(Sept. '74)	

Of these people, just over half say they find the advertising to be believable -- but a rather large 45% are not convinced.

#### Believability of Oil Company Advertising

Very believable	10%	} 55%
Fairly believable	45	
Not very believable	26	} 45
Not at all believable	19	
(Oct. '74)		

The public is not in the mood to be a friend of the oil companies, particularly as the companies report increasing profit levels. Eighty percent of the public supports the notion that the federal government should put a limit on oil company profits, even when the no-limit alternative is phrased in terms of encouraging increased production.

#### Oil Company Profits

Put federal limit on oil company profits	80%
Allow oil companies to make all they can to encourage increased production	13
No opinion	7
(Sept. '74)	

It has been difficult for the public to accept the energy shortage as something that has not been, at least in part, contrived by business. Although more people today than several months ago believe the shortage is real, over half the public believe that the situation has been and still is contrived to at least some extent.

Of course, even many experts -- self-appointed and otherwise -- have publicly disagreed as to the seriousness and reality of the shortage. Ralph Nader, on the one hand, has been saying all along that we are drowning in a sea

of oil. On the other hand, an official of the American Association for the Advancement of Science has said, "America's energy problems are real, serious, and enduring despite the popular belief that the difficulties vanished with the end of the Arab oil boycott."

Believe energy shortage is ...

	April <u>'74</u>	Sept. <u>'74</u>
Real	25%	35%
Contrived	43	38
Some of both	23	20

Whether they believe it real or contrived, very few people expect the energy shortage to become less severe over the next few months, and almost four persons in ten are anticipating an increased problem.

Expect energy shortage over  
next few months to become ...

	March <u>'74</u>	June <u>'74</u>	Oct. <u>'74</u>
More severe	11%	27%	38%
Less severe	36	11	5
Same as now	48	59	52

There are a whole host of possible actions or policies that might be undertaken to help solve or lessen the energy shortage. We have measured public favor for or opposition to 24 such actions or policies.

Actions that relate to providing increased information enabling the consumer to conserve energy usage are favored by a very large proportion of the public.

Consumer Information

	<u>Favor</u>	<u>Oppose</u>
Oil industry inform on energy conservation	88%	8
Appliance manufacturers inform on electric usage	86%	11
Federal gov't inform on energy conservation	84%	14

(Oct. '74)

Actions that would have a direct effect on consumer behavior receive mixed reaction. Actions that are not particularly limiting of behavior, such as requiring people to save materials for recycling and having a nationwide speed limit of 55 mph, receive the approval of the large majority of the public. However, actions that would have a more drastic effect on people's behavior, such as the regulation of energy use and the rationing of gasoline, do not go over particularly well.

#### Actions Affecting Consumer Behavior

	<u>Favor</u>	<u>Oppose</u>
Require saving materials for recycling	88%	9
Nationwide maximum speed of 55 mph	76%	23
Regulate energy use	42%	49
Ration gasoline	16%	80

(Oct. '74)

By and large, large majorities of the public approve of efforts to develop more energy sources, such as greater involvement of the federal government in developing new sources of energy, increased off-shore drilling, the building of more nuclear power plants, and government involvement in the increased exploration for oil. On the issue of tax incentives for oil companies to encourage increased production and exploration, public opinion is divided, with somewhat more opposition than favor.

#### Energy Development

	<u>Favor</u>	<u>Oppose</u>
Federal gov't develop new sources of energy	82%	10
More offshore drilling	74%	16
More nuclear power plants	66%	24
Federal gov't explore for oil	65%	25
Tax incentives to increase production and exploration	40%	50

(Oct. '74)

With respect to the easing of pollution controls to help alleviate the energy shortage, public opinion is rather divided:

On balance, the public approves of the removal of off-shore drilling restrictions.

Public opinion is evenly divided on the suspension of auto emission controls.

There is more opposition than favor, however, for the removal of strip mining restrictions and, particularly, for the removal of oil refinery pollution controls.

#### Pollution Controls

	<u>Favor</u>	<u>Oppose</u>
Remove offshore drilling restrictions	47%	40
Suspend auto emission controls	45%	44
Remove strip mining restrictions	32%	46
Remove oil refinery pollution controls	25%	67

(Oct. '74)

People also have mixed emotions about using taxes to control energy usage. There is slightly more opposition than favor to two suggested approaches -- namely, taxing large cars that get low mileage and taxing auto features (such as air conditioners) that decrease gas mileage.

#### Taxes To Control Usage

	<u>Favor</u>	<u>Oppose</u>
Tax on auto features that decrease gas mileage	45%	50
Tax on large cars that get low gas mileage	43%	49

(Oct. '74)

Opposition is great to actions that use price increases as a means of controlling energy usage. People overwhelmingly reject higher electric rates for peak-hour usage and increased gasoline prices to reduce consumption.

#### Price To Control Usage

	<u>Favor</u>	<u>Oppose</u>
Higher electric rates during peak hours	28%	68
Increase gas prices to reduce consumption	19%	77

(Oct. '74)

There is considerable support for increased federal spending on mass transportation and majority support for increasing the amount of foreign oil imports. There is more favor than opposition for the federal government becoming involved in the regulation of energy production, although this favor is barely at the majority level. And the notion of all-year Daylight Savings Time has never particularly stirred the public.

#### Miscellaneous Actions

	<u>Favor</u>	<u>Oppose</u>
Federal spending on mass transit	79%	17
More foreign oil imports	54%	38
Federal gov't regulate energy production	49%	38
All year D.S.T.	41%	53

(Oct. '74)

Most Americans -- 87% -- claim to be making some effort to save energy.

Effort made to save energy ...

Great deal	36%
Fair amount	51
Very little	10
None	3

(Oct. '74)

As shown here, one of the reasons that many people are making an effort to save energy is because they believe individual efforts do have an impact on total consumption. People also are making additional efforts today to save energy because of the rising prices.

Impact of personal conservation on total consumption

Great deal	34%
Fair amount	38
Very little	19
None	4
Don't know	5

(Oct. '74)

In our most recent survey, people report doing the following things to save energy:

Two-thirds say they are using fewer lights.

About half say they are using appliances less often. During the summer months there was a sharper rise in this kind of behavior due primarily to claimed reductions in air conditioning usage.



And about one person in eight claims to have postponed or cancelled an appliance purchase because of the energy shortage.

Because of energy shortage ...

Use fewer lights	67%
Use appliances less often	48%
Postponed or canceled appliance purchase	13%

(Oct. '74)

Over eight people in ten say they are driving slower in order to conserve energy. As mentioned before, the primary component in this behavior is now cost rather than availability.

Driving Slower --

Yes	84%
No	16

(Oct. '74)

Changes in shopping behavior may be one of the permanent results of the energy shortage. What we have seen since January is a fairly significant proportion of the public claiming to have changed some of their shopping patterns, and the proportion claiming to have changed has held up -- even increased -- over the months.

Over seven persons in ten say that they shop more at stores close to home or work than they did before. Two-thirds say that they shop at fewer stores or fewer locations in order to conserve energy. Six consumers in ten say they are shopping less often. And just over half report a reduction in window shopping.

Shopping Changes Due to Energy Shortage

	<u>June</u>	<u>Latest</u>
Shopping more at stores close to home or work	63%	72%
Shopping at fewer stores or locations	55%	68%
Shopping less often	53%	60%
Cutting down on "window shopping"	50%	54%

(Sept. '74)

One aspect of the energy shortage has been the rather sharp rise in electric rates over the past year. As shown here, large majorities of the public throughout the country testify to these increases. And very few people expect the situation to change in the near future.

#### Electric rates in past year ...

	Total	East	Mid-west	South	West
Increased	74%	72%	66%	79%	78%
Decreased	1	1	0	1	1
No change	12	12	16	10	8
Don't know	8	6	13	6	8

(Oct. '74)

Over half the public reports that they are using less electricity today because of the cost.

#### Using less electricity because of cost ...

Yes	55%
No	40
Don't know	5

(Oct. '74)

#### Implications

Over the long run in a market economy, price and availability are inextricably intertwined. The impact of this relationship is now being felt in a direct way by the public; and, as a result, their attitudes and behavior appear to be undergoing a slow but long-term shift. The potential implications of this shift are widespread and far from totally clear at this time. Among some of the areas that may well be impacted are these:

The generation of electricity -- There is the potential for increased movement to public ownership of electric utilities as a result of rapidly rising costs. Just recently in Massena, New York, it was voted upon to buy out Niagara Mohawk.

The pricing of electricity -- Over the long term, it may be necessary to price electricity in a manner similar to toll telephone charges -- i.e., higher prices during periods of peak usage in an attempt to smooth demand and lower costs.

Oil and automobile taxes -- There has been an increase in public demand for the elimination of the oil depletion allowance that seems to be tied to public dissatisfaction with the level of oil company profits. With respect to taxes on larger automobiles, there is considerable public opposition at the present time. However, the long-term potential for public support in this area should not be written off, as the desire for larger cars certainly will be impacted by their upkeep costs.

The warm, snug house -- The sharp rises in the price of fuel oil have made home heating a much higher proportion of total housing costs. In the Northeast, for instance, heating oil rises have eclipsed real estate tax rises as a prime component in the inflation of housing costs. The result is a substantial growth in the desire for "warm, snug houses" with resulting impact on the markets for everything from draperies and insulation to the houses themselves. Attached houses, in particular, which have been relatively successful in the last few years in reaction to land costs, are relatively much cheaper to heat and presumably more desirable.

Mass transit -- There is growing public support for increased federal funds to be spent in the area of mass transit. Public demand for and usage of mass transit should be followed closely because of the obvious impacts that increased mass transit demand would have on a large number of policies and markets.

The credibility of business -- In today's adverse business climate, where the credibility of most major institutions is at a low point, the public's growing skepticism about the actual nature of the energy crisis and the role of the oil industry in it could create a backlash against all business. In fact, the questions raised by the energy crisis may well intensify the public's growing criticism of what it considers to be the concentrated economic power, excessive profits, and impersonality of large companies. The explosive quality of the crisis is indicated by the way the public feels that, either voluntarily or under pressure, it's bearing the brunt of the shortage while those in power have not fully carried out their responsibilities to alleviate the situation.

PERCEPTIONS OF CHILDREN'S TELEVISION ADVERTISING:  
AN EMPIRICAL INVESTIGATION OF THE BELIEFS AND ATTITUDES  
OF CONSUMER, INDUSTRY, AND GOVERNMENT RESPONDENTS

James D. Culley<sup>1</sup>  
University of Delaware

This paper summarizes the findings of a study investigating the beliefs and attitudes of six key respondent groups regarding issues surrounding television advertising and children. The six groups included in the study are: spokesmen for Action for Children's Television (ACT); the presidents and top executive officers of advertising agencies creating and producing children's television commercials; top executives in companies advertising heavily on children's television programs; members of the Federal Trade Commission, the Federal Communication Commission, and key members of Congress; college students; and adult members of the general public. The paper reports respondent attitudes towards (1) increased regulation, (2) self-regulation and government regulation, (3) the techniques used in children's television commercials, (4) the effects of television commercials on children, and (5) the future of children's television advertising.

It has been said that much of the current controversy regarding television advertising and its effect on children is the result of ineffective communication between broadcasters, researchers, and critics of television broadcasting (Ward, 1972). Each of these interest groups pursues the subject with markedly different assumptions concerning the issues and with erroneous or incomplete perceptions of how the other parties involved view the same issues. The result is a series of "dialogues that never really happen" (Bauer & Greyser, 1967).

Although there exists a large number of studies focusing on the general public's attitudes towards advertising as a marketing tool, in only a few cases has an attempt been made to deal with areas of specific concern to those interested in children's television advertising. Moreover, a survey of the literature revealed practically no attitude studies of the key people involved in the creation, production, research, and evaluation of advertising directed at children.<sup>2</sup> Yet, these people are among the best sources of information on the various issues involved, the ones most likely to influence and to be influenced by policy decisions in the area, and the expert opinion leaders for much of the general public on the subject.

The purpose of this research study was to present an objective examination of the attitudes and co-orientation ability of seven key respondent groups towards the major issues surrounding the subject of children's television advertising. The seven groups included in the study were:

Action for Children's Television Spokesmen (ACT). This Boston based

consumer group has had considerable success in petitioning the Federal Communication Commission and the Federal Trade Commission for changes in present policies regarding children's television advertising. ACT has also been successful in raising money for research on the subject of children's television and children's television advertising and in influencing public opinion on the major issues involved with children's television advertising. A judgment sample of sixty ACT spokesmen was selected for the survey by the founders of ACT; eighty-five percent of the sample completed the survey.

Advertising Agency Executives. Much of the criticism of children's television advertising has to do with the techniques and the content of commercials aimed at children. In this study, 107 presidents and top executive officers of twenty-four advertising agencies that create, produce, and buy broadcast time on children's television programs were sent copies of the survey. Sixty-six percent completed the questionnaire. In gathering preliminary data for the study, personal interviews were also conducted with the top executives at ten of the twenty-four agencies included in the final survey.

Top Advertisers. There is a great variety of products advertised on children's television, but most fall into one of four categories: toys, cereals, candies, and other food items. In this study, seventy-five copies of the final questionnaire were mailed to the presidents and top executive officers of thirty firms sponsoring shows, or frequently running commercials, on network children's television. Forty-five percent of the sample mailed back usable questionnaires.

Members of the Federal Trade Commission (FTC), the Federal Communication Commission (FCC), and Key Members of Congress. A judgment sample of forty-nine congressmen was included in the government sample. These congressmen were selected because of their expressed interest in and voting record on consumer causes. In addition, fifteen commissioners and top staff officers of the FTC and the FCC were included in the sample. Sixty-six percent of the government sample responded to the survey, but only thirty-four percent of the sample sent back completed questionnaires. (Note: special care should be taken in interpreting the responses of the government sample for two reasons. One, only those congressmen considered "consumer oriented" by their colleagues were included in the sample. It was felt that these congressmen would be most likely to draft legislation in the area and be opinion leaders for others on the subject of children's television. Two, only a small percentage of the respondents in this sample sent back usable questionnaires.)

Members of the Network Review Boards. Major gatekeepers for all commercials destined for airing on network television are the network review or continuity boards. Though the number of people on these boards is small, the board members hold a major position in determining what is or is not acceptable for showing on the networks. Thirteen review board members from the three major networks were mailed copies of the survey, but only six responded. Due to the small size of the sample group and the low response rate, no statistical analyses were attempted.

College Students and the General Public. Several months after the first five respondent groups were surveyed by the author, Professors Frank N. Pierce and Leonard J. Hooper of the University of Florida surveyed random samples of Gainesville, Florida townspeople and college students using

portions of the author's questionnaire.

A mail questionnaire was used as the major research instrument for the initial survey. Personal interviews by trained student interviewers were used to gather data for the Pierce-Hooper survey.

The initial survey was divided into two major sections. The first section contained twenty-nine Likert-type attitudinal items designed to test the variance within and mean difference between the responses of the surveyed groups on major issues regarding children's television advertising. The second section of the questionnaire was designed to measure three relationships between the various respondent groups: the amount of cognitive overlap, or similarity in attitudes of the respondents; the perceived cognitive overlap, or extent to which each respondent thought his beliefs were the same as other participants completing the survey; and the accuracy of the respondents in estimating the position of other respondents on issues involving children's television advertising.

Pierce and Hooper designed two similar questionnaires for their study of student and general public attitudes towards television advertising and children. Both questionnaires contained over seventy items. Over one-third of the items were demographic in nature such as how long the respondent lived at his present address, the respondent's age, political party preference etc. Eleven items centered on the respondent's attitude toward advertising as an element in the marketing mix. The remaining twenty-nine items were taken from the first section of the author's study.

Six general hypotheses and twenty-two research hypotheses were included in the author's initial study. All the general hypotheses and seventeen of the research hypotheses concerned the following topics: (1) the need for regulation of children's television advertising; (2) the method of regulating children's television advertising; (3) the effects of television commercials on children; (4) the techniques used in commercials aired on children's television; (5) the products advertised on children's television; and (6) major proposals regarding the future of children's television advertising. Five research hypotheses concerned the ability of the individuals involved with children's television advertising to interact effectively on the issues involved.

Two mailings of the initial survey were made. The first mailing was made June 12, 1973. The follow-up mailing, sent to all respondents failing or refusing to complete the initial questionnaire, was made one month later. The Pierce-Hooper interviews were conducted in late April and early May, 1974.

T-tests, analysis of variance tests, and Duncan range tests were the basic statistical tools used in the analyses.

### Major Findings

Although six major issues were covered in the survey, thirteen of the twenty-nine items in the survey loaded heavily on two dimensions: the regulation dimension and the advertising as a profession dimension. The study findings relating to these two dimensions and the findings relating to the

co-orientation section of the survey are the subject of the remainder of this paper.

### The Regulation Dimension

The issue of whether there should be more regulation of children's television commercials divided the respondents into two distinct groups. All of the ACT sample, over ninety percent of the government sample, and over sixty-five percent of the student and townspeople samples agreed that television advertising to children should be more regulated than it already is. Sixty-six percent of the agency sample and eighty-five percent of the advertiser sample disagreed with the item: "television advertising to children should be more regulated than it already is."

All six respondent groups agreed that special regulation of children's television advertising was required because of the nature of the viewing audience. There were significant differences, however, in the strength of their agreement. Over fifty percent of the Hooper-Pierce sample and over eighty percent of the ACT sample indicated "strong agreement" with the item. Less than ten percent of the industry samples indicated "strong agreement."

Item 1: Television advertising to children should be more regulated than it already is.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean <sup>3</sup>
Students	31.7%	37.3%	10.3%	14.6%	6.1%	445	3.7
Townspeople	49.0	27.7	12.1	9.0	2.2	455	4.0
Agency	4.3	18.6	11.4	44.3	21.4	34	2.4
Advertiser	2.9	5.9	5.9	58.8	26.5	71	2.0
ACT	90.9	9.1	--	--	--	51	4.9
Government	40.9	50.0	--	9.1	--	22	4.2

Item 15: Children's television advertising requires special regulation because of the nature of the viewing audience.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean
Students	49.0%	41.3%	3.1%	4.9%	1.6%	445	4.3
Townspeople	56.3	33.2	5.5	3.3	1.8	455	4.4
Agency	8.6	52.9	7.1	20.0	11.4	34	2.9
Advertiser	8.8	47.1	11.8	23.5	8.8	71	2.8
ACT	84.1	13.6	2.3	--	--	51	4.8
Government	27.3	59.1	4.5	9.1	--	22	4.1

Who should regulate? Items 4, 16, and 18 on the survey referred specifically to the issue of how to regulate television advertising directed at children.

Over sixty percent of the respondents in the industry samples favored self-regulation. A majority of the student and townspeople samples agreed. The government sample favored self-regulation but over sixty percent of the sample felt the trade association guidelines in use today have done little to improve the quality of children's television advertising.

Over seventy percent of the agency sample and all of the advertiser sample felt the trade association guidelines have helped improve the quality of commercials aimed at children.

Seventy percent of the ACT sample felt that commercials to children should be regulated by the government instead of by advertisers.

Item 4: Commercials to children should be regulated by advertisers themselves.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean
Students	26.5%	36.0%	10.5%	20.7%	6.3%	445	3.5
Townspeople	27.5	26.2	12.2	22.9	11.2	455	3.4
Agency	20.6	42.6	5.9	22.1	8.8	34	3.4
Advertiser	11.8	55.9	11.7	20.6	--	71	3.6
ACT	4.5	9.1	4.6	13.6	68.2	51	1.7
Government	4.5	18.2	4.6	54.5	18.2	22	2.4

Item 16: The trade association guidelines in use today have done little to improve the quality of children's television advertising.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean
Students	12.1%	25.4%	29.9%	28.3%	4.3%	445	3.1
Townspeople	20.7	23.5	30.1	20.2	5.5	455	3.3
Agency	--	15.7	11.4	40.0	32.9	34	2.1
Advertiser	--	--	--	67.6	32.4	71	1.7
ACT	61.4	36.4	2.2	--	--	51	4.6
Government	4.5	59.1	22.8	13.6	--	22	3.6

Item 18: Commercials to children should be regulated by the government.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean
Students	8.8%	21.6%	6.6%	30.6%	32.4%	445	2.0
Townspeople	10.3	20.2	10.2	25.7	33.6	455	2.5
Agency	2.9	8.6	4.2	30.0	54.3	34	1.8
Advertiser	--	11.8	2.9	44.1	41.2	71	1.9
ACT	38.6	31.8	11.5	13.6	4.5	51	3.9
Government	18.2	31.8	31.9	13.6	4.5	22	3.5

The matter of numbers. Over half of all six respondent groups felt there are too many commercials on shows children watch. There were significant differences, however, in the strength of their agreement. For example, ninety-three percent of the ACT respondents and fifty percent of the government respondents marked "Strongly Agree" on the item. Only twelve percent of the agency sample and none of the advertiser sample marked the "Strongly Agree" category.

Eighty-nine percent of the ACT respondents favored banning all commercials from children's television programs. Less than half the student, townspeople, agency, advertiser, and government samples agreed.



Item 6: There are too many commercials on shows children watch.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean
Students	29.0%	36.2%	15.9%	15.5%	3.4%	445	3.7
Townspeople	43.1	27.7	14.5	10.5	4.2	455	4.0
Agency	11.8	38.2	5.9	33.8	10.3	34	3.1
Advertiser	--	50.0	9.4	37.5	3.1	71	3.1
ACT	93.2	6.8	--	--	--	51	4.9
Government	50.0	45.5	4.5	--	--	22	4.5

Item 22: Advertising on children's television programs should be banned completely.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean
Students	7.9%	14.4%	5.4%	38.4%	33.9%	445	2.2
Townspeople	17.6	14.7	5.3	35.8	26.6	455	2.6
Agency	1.4	1.4	2.9	24.3	70.0	34	1.4
Advertiser	2.9	--	--	5.9	91.2	71	1.2
ACT	72.7	15.9	11.4	--	--	51	4.6
Government	13.6	22.7	31.9	31.8	--	22	3.2

#### The Advertiser and His Product: Are they Good?

Six of the 29 items in the study seemed to center on the respondents' reactions toward advertisers as basically honest or laudable persons, and toward their product (the television commercial directed at children) as an honorable and ethical business practice. Agreement with any of these statements, whether held strongly or lightly, is taken to mean that the person interviewed identifies more closely with the stance of the television advertiser than he does with the stance of those organizations who oppose current practices. The respondent who agrees for whatever reasons is more a supporter of the status quo and less fearful of the results of television commercial consumption by children than is his disagreeing counterpart.

The student and townspeople samples saw fit to reject a pro-advertiser position rather strongly in four of the six statements presented to them. They did agree with that stand twice, but each time by slight margins and once not even half of those who held opinions were in agreement. Scores ranged from a low of 19 percent agreement on the least favorably regarded statement to a high of 53 percent on the most favorably regarded one. The average amount of agreement--35 percent over the half-dozen statements--was far outweighed by the 56 percent of the populace who disagreed. Nine percent were neutral.

The question of truth. Not one respondent of the ACT sample, less than ten percent of the government sample, and less than twenty percent of the student and townspeople samples felt children's television commercials present true pictures of the products advertised. Over sixty-five percent of the advertiser sample and over seventy-five percent of the agency sample felt that they do.

Over eighty percent of the industry respondents felt that advertisers make a sincere effort to present their products truthfully. Less than half the student, townspeople, ACT, and government samples agreed.

Item 8: Most children's television commercials present a true picture of the product advertised.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean
Students	2.5%	16.0%	4.5%	41.3%	35.7%	445	2.1
Townspeople	5.1	14.9	9.0	34.3	36.7	455	2.2
Agency	10.0	57.1	18.6	11.4	2.9	71	3.6
Advertiser	14.7	61.8	14.7	8.8	--	34	3.8
ACT	--	--	--	25.0	75.0	51	1.3
Government	--	9.1	9.1	59.1	22.7	22	2.0

Item 19: Most advertisers on children's television make a sincere effort to present their product truthfully.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean
Students	1.3%	27.9%	7.2%	39.6%	24.0%	445	2.4
Townspeople	5.9	29.9	12.6	30.5	21.1	455	2.7
Agency	24.3	60.0	10.0	5.7	--	71	4.0
Advertiser	32.4	61.8	5.8	--	--	34	4.3
ACT	--	--	4.5	43.2	52.3	51	1.5
Government	--	13.6	27.3	50.0	9.1	22	2.5

The question of taste. Fifty-three percent of the Pierce-Hooper survey respondents agreed that most television commercials aimed at children are in good taste. Only one-third of the student and townspeople samples disagreed. Over seventy percent of both industry samples also agreed with the item.

At the other extreme on the taste dimension stood the ACT and government samples. Not one of the ACT respondents and only fourteen percent of the government respondents felt that commercials aimed at children are in good taste.

Item 11: Television commercials aimed at children are usually in good taste.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean
Students	10.3%	45.4%	11.3%	21.8%	11.2%	445	3.2
Townspeople	11.2	38.7	14.5	20.4	15.2	455	3.1
Agency	14.5	59.4	11.6	14.5	--	34	3.7
Advertiser	14.7	64.7	11.8	8.8	--	71	3.9
ACT	--	--	2.3	43.2	54.5	51	1.5
Government	--	13.6	9.2	54.5	22.7	22	2.1

Advertising as a marketing tool. Sixty-five percent of the agency respondents and seventy-one percent of the advertiser respondents felt that commercials help develop a child's ability to make good consumer decisions. Less than five percent of the ACT sample and less than twenty-five percent of the townspeople, student, and government samples agreed.

Item 2: Advertising helps develop a child's ability to make good consumer decisions.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean
Students	1.8%	17.5%	6.3%	41.1%	33.3%	445	2.1
Townspople	7.3	20.2	8.9	28.4	35.2	455	2.4
Agency	10.3	54.4	17.7	14.7	2.9	71	3.5
Advertiser	5.9	64.7	20.7	8.8	--	34	3.7
ACT	2.3	--	2.2	20.5	75.0	51	1.4
Government	4.5	18.2	9.1	45.5	22.7	22	2.3

Advertisers as professionals. Over seventy percent of the student, townspeople, ACT, and government samples agreed with the statement: "Most advertisers on children's television are not really concerned about kids; they just want to sell their products."

Less than half the townspeople, student, ACT, and government samples felt advertisers were trying their best to provide what the public wants.

Item 26: Most advertisers are good people trying their best to provide what the public wants.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean
Students	3.6%	25.4%	6.7%	35.1%	29.2%	445	2.4
Townspople	10.1	31.4	7.5	28.6	22.4	455	2.8
Agency	27.9	50.0	14.7	7.4	--	34	4.0
Advertiser	38.2	50.0	11.8	--	--	71	4.3
ACT	--	2.3	6.4	36.4	54.9	51	1.6
Government	--	31.8	31.8	36.4	--	22	3.0

Item 27: Most advertisers on children's television aren't really concerned about kids, they just want to sell their product.

	Strongly Agree	Somewhat Agree	Uncertain/ No Answer	Somewhat Disagree	Strongly Disagree	# in Group	Mean
Students	34.6%	40.2%	5.7%	18.2%	1.3%	445	3.9
Townspople	37.1	34.3	6.4	17.1	5.1	455	3.8
Agency	4.3	20.3	7.3	47.8	20.3	34	2.4
Advertiser	--	2.9	14.8	44.1	38.2	71	1.8
ACT	75.0	22.7	--	--	2.3	51	4.7
Government	4.5	72.7	4.6	18.2	--	22	3.6

#### The Co-orientation Abilities of the ACT, Agency, Advertiser, and Government samples

The initial mail survey was divided into 2 sections. The first section centered on the twenty-nine attitudinal items discussed up to now. The second section of the survey dealt specifically with the co-orientation abilities of the ACT, Agency, Advertiser, and Government samples.<sup>4</sup>

The first relationship measured in this section concerned the amount of cognitive overlap, or similarity in attitudes of the various respondent groups.

On twenty-three of the twenty-nine items in the survey, the mean response of the two industry samples differed significantly from that of the ACT respondents. The items in which the direction of the mean scores of the industry and ACT respondent groups were the same are:

- : There are too many commercials on shows children watch. (The agency, advertiser, and ACT respondents agreed.)
- : Commercials often persuade children to want things they don't really need. (The agency and ACT respondents agreed.)
- : Performers should be allowed to sell products on children's television shows. (The agency, advertiser, and ACT respondents disagreed.)
- : Simulcasts (permitting 2 or more networks to run the same program at the same time) would help improve the quality of children's television programming. (The agency, advertiser, and ACT respondents disagreed. However, nearly seventy percent of the ACT respondents marked "uncertain" on this item. It is doubtful whether the ACT respondents had heard of the simulcast proposal before.)
- : Children's television advertising requires special regulation because of the nature of the viewing audience. (The agency, advertiser, and ACT respondents agreed.)
- : It is up to the parents to regulate children's television viewing behavior. (The agency, advertiser, and ACT respondents agreed.)

The advertiser and agency respondents differed significantly in their attitudes on only four of the twenty-nine Likert-type items. In each case, the direction of agreement was the same: only the expressed strength of the responses differed. For example, on the statement, "Most advertisers on children's television aren't really concerned about kids, they just want to sell their products," both the agency and advertiser respondents expressed disagreement. However, while thirty-eight percent of the respondents from the advertiser sample expressed strong disagreement with the item, only twenty percent of the agency respondents did. A two-tail T-test comparing the sample means of the two groups did not support the hypothesis that the samples were drawn from the same population.

The other three items in which tests of the differences between the mean scores of the industry respondents were significant are:

- : "Bunching" commercials before or after a program would significantly lessen the impact of the advertiser's message. (The respondents in the advertiser sample expressed stronger agreement than the respondents in the agency sample.)
- : The trade association guidelines in use today have done little to improve the quality of children's television advertising. (The respondents in the advertiser sample expressed stronger disagreement than the respondents in the agency sample.)
- : Television commercials lead to an increase in parent-child conflict. (The respondents in the advertiser sample expressed stronger disagreement than the respondents in the agency sample.)

The second relationship measured in the co-orientation section of the survey dealt with the accuracy of the respondents in estimating the position of other respondents on the items included in the co-orientation section of

the survey. Five different data sets were used in testing respondent accuracy: the ACT respondents' estimates of the agency position, the agency respondents' estimates of the ACT position, the advertiser respondents' estimates of the ACT position, and the government respondents' estimates of the ACT position.

Only the data from the ACT respondents' estimates of the agency position supported the hypothesis that the accuracy of each respondent group will be relatively low in estimating the position of other groups on issues relating to children's television advertising.

The third relationship measured in the co-orientation section of the survey dealt with the perceived cognitive overlap or extent to which each respondent thought his beliefs were the same as other participants completing the survey. The actual position of the agency sample differed significantly from its perception of the ACT position on all nine co-orientation items. But there was no significant difference between the actual agency position and its perception of the network position on six of the nine items.

The actual position of the advertiser respondents differed significantly from their estimate of the ACT respondents' position on all nine items. But there was no significant difference between the agency respondents' position and their perception of the network respondents' position on six of the nine co-orientation items.

The position of the ACT respondents differed significantly from their estimate of the network respondents' and agency respondents' positions on all nine co-orientation items. The position of the government respondents differed significantly from their estimate of the ACT respondents' position on four of the nine co-orientation items. The position of the government respondents differed significantly from their estimate of the network respondents' and agency respondents' position on eight of the nine co-orientation items.

### Implications of the Study

In recent years, government and business spokesmen have advocated a dialogue between key government, industry, and consumer groups to reduce friction and advance the general good. Yet, such a dialogue never happens. Rather, what passes for dialogue in form is only a sequence of monologues in fact, wherein each spokesman merely grants "equal time" to others and pretends to listen while actually preparing his own set of comments (Bauer & Greyser, 1967; Ward, 1972).

The findings of this study indicate that the lack of effective interaction between government, industry, and consumer spokesmen, at least regarding children's television advertising, is not due to a lack of understanding. The various groups included in the study do understand each other's positions on most issues remarkably well.

The findings indicate the critical need for industry spokesmen to establish an effective dialogue with government representatives and consumer spokesmen, particularly spokesmen for Action for Children's Television, if they hope to continue operating with the relative freedom they now enjoy. On almost every issue in the survey, the government respondents were on the side of the ACT respondents and not the side of the industry--a fact that has

serious implications regarding present public opinion and potential legislation on the subject.

Although the study data supported most of the research hypotheses, the significant differences in attitude between the surveyed groups and the small amount of variance in attitude within the surveyed groups is probably a more significant finding than the fact that the hypotheses were supported. The difference in attitudes between the industry respondents and the ACT and government respondents is so large that no publicity campaign or goodwill effort on the part of the industry is likely to have much effect on either group--at least not in the immediate future.

#### FOOTNOTES

1. James D. Culley is an Assistant Professor of Business Administration at the University of Delaware. The majority of the work described in this paper was completed while the author was a doctoral student at Michigan State University.
2. When children's television is referred to in this report, it means regularly scheduled network programs for which children make up the largest percentage of the viewing audience. "Children" designates any individual aged 12 or younger. Most network programs on Saturday and Sunday from 8:00 a.m. to 1:30 p.m. and The Captain Kangaroo Show weekdays on CBS would, therefore, be considered children's television shows.
3. Scores of 5, 4, 3, 2, and 1 were assigned to the 5 response categories from Strongly Agree to Strongly Disagree. Mean scores therefore could range from 1.0 to 5.0. The higher the mean score, the more group disagreement with the item as stated.
4. Nine of the twenty-nine items used in the first section of the survey were repeated in the second section of the questionnaire. Respondents were asked in the co-orientation section to respond as if they were someone else. For example, those in the ACT sample were told to respond as if they were spokesmen for firms advertising on network children's television or spokesmen for the major network continuity boards.

#### REFERENCES

- Bauer, Raymond A., & Greyser, Stephen A. Thinking ahead: The dialogue that never happens. Harvard Business Review, 1967, 45, 2-12, 186-190.
- Ward, Scott. Kids TV-marketers on hot seat. Harvard Business Review, 1972, 50, 16-18ff.



CURRENT COSTS OF FAMILIES' PAYMENTS COMPARED TO ESTIMATES OF HEALTH PLANS<sup>1</sup>

Flora Williams<sup>2</sup>  
Purdue University

The current costs of families' fixed payments for medical care were compared to estimated fixed payments under proposed health plans. Families selected randomly from six metropolitan areas were paying an average of \$149 per year for total fixed medical costs compared to the estimated \$84 insurance cost under the Health Security Plan, \$150 under the Employee Health Insurance Plan, and \$600 under the Assisted Health Insurance Plan. Total fixed medical costs were composed of expenses for health insurance which averaged \$35, payroll deductions which averaged \$27, and fixed or regular health costs which averaged \$87. Families at every income level had fixed expenses greater than they would have under the Health Security Plan. Families under \$7500 income were paying more than they would under proposed EHIP costs but families over \$7500 were paying less.

Various plans of paying health expenses are being studied to assure good health care for all Americans. Senator Kennedy (Congressional Record, 1973), when discussing the crisis in rising health care costs, said the answer is "not to cut back on benefits, to raise insurance premiums even more, or to simply offer more insurance to more Americans. The answer is to reform our health care system and bring these costs under control."

The proposals currently before Congress outline in detail the structure, benefits, reimbursement policy, regulation and administration, financing, and special provisions. Not reported is the immediate impact on the consumer's budget allocations. Consumer behavior would probably be adjusted as fixed commitments are established by whichever plan is adopted. Flexibility in expenditures for health care might be decreased even if the coverage is better and the cost pressures of irregular medical expenses are reduced. The immediate impact of any increased fixed commitments for health maintenance might be felt before most consumers realize the improved health benefits.

One question to be asked when studying various health plans for predicting the immediate impact on consumers' financial behavior is: Would one plan more than another more closely match what families themselves pay currently in terms of fixed costs? This question looks only at the impact of changes in fixed commitments regardless of additional benefits or additional expenses for deductibles or expenses not covered by the plans.

What are the fixed expenses families have had for their medical costs? Which of several selected variables influenced the amount of insurance payments by families? What would be the possible changes in amount of money allocated by families for fixed medical costs when the new plan would be adopted nationwide? Answers to these questions might influence the initial reaction of families to any change.



Fixed medical costs in this study were defined as the total money spent by families for health insurance, payroll deductions for health insurance, and health expenses reported by the families as fixed. These costs could be compared to aggregated estimates of fixed costs under Senator Kennedy's "Health Security Plan" and the Administration's "Comprehensive Health Insurance Plan." Not examined were extent of coverage, employer's contributions to insurance coverage, deductibles, or health expenses that would be irregular or not fixed. Therefore, such provisions in the health plans are not discussed in depth or included in the cost comparisons for this report.

The proposed Health Security Act of 1973 would cover every resident with health insurance regardless of residence, work status, income, size of family, or history of medical problems. Payment for the coverage would be based on a person's ability to pay. The Health Security Plan converts

The existing Medicare hospital insurance payroll taxes into Health Security taxes, and raise the rates to 1 percent on employees and 3.5 percent on employers. ...raises the wage base for the employer tax from \$12,000 (as present law provides after 1973) to \$15,000, with subsequent further increase if wage levels rise, so that the Health Security Wage base will always be 125 percent of the Social Security wage base. ...adds a new 1 percent Health Security tax on unearned income (unless such income is less than \$400 a year), subject to the same maximum on taxable income as is applicable to the employee and self-employment taxes. Taxable unearned income is adjusted gross income up to the stated maximum, minus wages and self-employment income already taxed for Health Security purposes (excluding certain items of income specifically excluded from the other taxes and excluding \$3,000 in unearned income for persons over age 60) (Congressional Record, January 31, 1973, p. 30).

The two structures of the Comprehensive Health Insurance Plan (The White House, 1974) are the (a) Employee Health Insurance Plan (EHIP) whereby all employers would be required to offer the basic insurance plan or Health Maintenance Organization coverage while the plan would also be available to self-employed and non-working families; and (b) Assisted Health Insurance Plan (AHIP) whereby families would enroll who are not covered under EHIP, are below \$5,000 income, are non-working or high risk with income between \$5,000 and \$7,500, or work in unusually high risk insurance type jobs.

Employers would contribute 65 percent of the premium expenses for covered employees with Federal Government subsidy if an employer's payroll rises by more than 3 percent due to required contributions to coverage. Eventually the employer would pay 75 percent of premium costs and employees the remaining 25 percent. "Premiums for employer groups of 51 or more employees and other families and groups being offered EHIP would be negotiated between employer and other groups and the insurance carrier" (The White House, 1974, p.4). The average family cost for health insurance premiums would be about \$600 per year with the employee paying approximately \$150 per year with the employer paying approximately \$450 for each employee (Congressional Record, February 6, 1974, S1341).

The premiums, deductibles, coinsurance, and maximum liability would be related to income under the AHIP. Working families with income below \$5,000 would pay no premiums at all.

## 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 interviewers. 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.

### 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 Bureau of the Census (Maps of Major Concentrations of Poverty, 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 1964-1967, 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 amount of payments for insurance. 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).

### Characteristics of the Sample

Characteristics of the families in this sample from metropolitan areas varied widely in their demographic characteristics. By ethnicity, 49 percent were white, 30 percent were black, 5 percent were oriental, 12 percent were Spanish-Americans, and 4 percent were of other ethnic origins. The mean total disposable income was \$8,434.51.

Characteristics of the sample families in 1970 may be compared to national figures in 1970 in order to judge application of findings to families other than those in the sample. The percentage of the families in the sample who had over \$10,000 income was 22, whereas in the nation as a whole it was 49 (Statistical Abstract, 1971, p. 322). 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 national figure for families below the poverty level was also 15 percent (Statistical Abstract,

1971, p. 323). 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. The national median was 12.2 (Statistical Abstract, 1971, p. 111).

### Data Analysis

Data from these six metropolitan areas were combined for analysis purposes. The number of families supplying useable 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.

The one-way analysis of variance technique was used to determine differences of means between fixed health expenditures by selected variables. The Pearson's product-moment correlation was used to determine the relationship of selected variables with fixed medical expenditures. Multiple regression analysis was used to measure the simultaneous effects of different variables and to indicate the most significance variables related to payment for health insurance. Unless indicated otherwise, the .05 level of significance was used to determine statistically significant differences and relationships.

### Selected Variable Definitions

Health insurance expenditures were reported by the families in answer to the questions, "Could you list the bills or expenses you are supposed to meet regularly? That is, the sort of things you feel are rather fixed, or that you are obliged to pay, or that you have promised to pay each week or every month or every so often, or that are taken out of a paycheck?" The amount for a given period was reported. The number of payments due the past 12 months, the dollar amount per year, and dollar amount that was a payroll deduction were recorded.

Although health insurance premiums buy various types and amounts of coverage, this study looked at only the annual premium payment by the family for that coverage. Employers may have paid all or part of health insurance premiums but this study looked only at the payments or contributions made by the families.

Health costs were similarly reported. The dollar amount considered as the medical bills or expenses the family was supposed to meet regularly, felt obligated to pay, or promised to pay were recorded as fixed expenditures. Examples of these were regular purchase of drugs such as insulin or periodic payments on a hospital or dental bill. The total amount per year was calculated after recording the number of payments due during the past 12 months.

Life insurance payments were reported in the same manner as health insurance payments and calculated on an annual basis. This variable was included as an indicator of insurance orientation.

Total disposable income was the flow of purchasing power received from investments, social security benefits, job-related benefits such as disability or unemployment insurance, armed services benefits, welfare benefits, legal arrangements, and earned income. Earned income did not include FICA or income taxes. Income was grouped to correspond to annual income groups in the schedule of cost estimates for the "Comprehensive Health Insurance Act of 1974."

Income index was computed by dividing the total family disposable income by its poverty threshold and multiplying by 100. The poverty threshold as measured by Orshansky (1969) attempts to take into account the composition and the number in the family and is based on the assumption that no more than a third of the family income is used for food. An income index of less than 100 indicates that the family's income is below its poverty threshold when this threshold is based on the economy food plan of the USDA, a plan designed for short term emergency situations.

In addition to the objective measure of income adequacy, the perceived adequacy of income was examined. Perceived adequacy was reported from the question, "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.

A financial problem of money for medical costs was identified when the respondent was asked, "Do you have the problem of not enough money for dentist, doctor, or medicine?" Responses were: (1) never, (2) seldom, (3) sometimes, and (4) often.

Ethnicity was observed by the interviewer and respondents were classified as white, black, oriental, Spanish-American or other. "Other" included Poly-nesi-ans, Indians, mixture, and unknown.

Education referred to the number of years completed in school. Education was reported for the respondent and for her husband.

Occupation was determined by asking the respondent to describe the kind of job held and what business or product was involved. The job contributing the most to the respondent's or her husband's total income was used to determine the occupation held. The scale with 29 levels was based on North-Hatt and Duncan's rankings (Reiss, 1961).

Disability was derived from the question, "Is anyone in this family sick all the time or disabled in any way?" Responses were: (1) not disabled, (2) not limited in any of the following ways, (3) able to work, keep house, go to school, or play, but limited in other activities, (4) able to work, etc., but limited in kind or amount of work, etc., and (5) not able to work, etc., at all.

## Findings

### Fixed Expenses Families Have Had for Medical Costs

Health Insurance Excluding Payroll Deductions. Eighty-three percent of all families did not report having fixed expenses for health insurance excluding payroll deductions, although the percent covered by insurance is not known for this sample. This large number of families with no insurance expenses brought the average amount for health insurance payments to a low of \$35.

The average for just those families who had any expenses for health insurance was \$212 (Table 1). The means did not differ significantly among families in various income groups. The maximum spent for insurance by any family was over \$600 for families in the three lowest income levels and just under

TABLE 1

Health insurance, fixed health costs and payroll deductions by income level.

Income level	Total families	Health Insurance <sup>a</sup>			Fixed Health Costs <sup>b</sup>			Payroll Deductions <sup>c</sup>		
		No.	Mean	Std. dev.	No.	Mean	Std. dev.	No.	Mean	Std. dev.
\$0 - 2,499	118	7	\$189	\$214	18	\$403	\$323	1	\$467	
2,500 thru 4,999	257	28	\$209	\$157	56	\$318	\$533	10	\$107	\$ 82
5,000 thru 7,499	357	56	\$214	\$142	122	\$286	\$274	63	\$184	\$120
7,500 thru 9,999	234	43	\$200	\$143	71	\$231	\$190	45	\$168	\$112
10,000 and above	273	72	\$218	\$113	75	\$417	\$553	63	\$205	\$110
Total	1239	206	\$212	\$137	342	\$315	\$395	182	\$185	\$116

<sup>a</sup> F ratio = .158, not significant beyond the .05 level.

<sup>b</sup> F ratio = 2.49, significant beyond the .05 level.

<sup>c</sup> F ratio = 3.48, significant beyond the .01 level.

\$500 for the two highest income levels. Families classified as low income, below \$5000, were 30.3 percent of the sample. About 90 percent of these low-income families were not paying any health insurance premiums themselves.

Families who had no fixed payments for health insurance were compared with families who had \$1 to \$199, those who had \$200 to \$299, and those who had \$300 or more expenses by the analysis of variance test. Significant differences ( $p = < .01$ ) among groups occurred for mean education of husband, total disposable income and income index. The average years of schooling for the no-expense group was 10.49 compared to 11.14, 12.43 and 11.42 for other groups, respectively. One might argue the explanation that the low educated husbands would be employed in jobs providing some health insurance coverage. The argument probably would not hold, however, because the mean occupational level did not differ significantly among the expense groups and the Pearson-product moment correlation value ( $r = -.09$ ) for husband's occupational level and expenses for health insurance was negative ( $p < .05$ ).

The average total disposable income for the no-payment group was \$6,934 compared to \$8,779, \$9,384, \$8,040 for the other groups, respectively, by level of health insurance payments. The average income index for the no-expense group was 161 compared to 218, 262 and 247 for the other groups, respectively.

Payroll Deductions. Payroll deductions occurred for only 15 percent of the families. For all families, the average was \$28.24. For all those families who had payroll deductions the average amount was \$185 (Table 1). Means of payroll deductions differed by income level. The lowest mean was for families in the \$2,500 through \$4,999 income group.

Payroll deductions for just the husbands averaged \$191 (Table 2). Two thirds of the husbands had deducted this amount plus or minus \$114. The means for various occupational groups did not differ significantly. However, the percentage having any deductions for health insurance did vary considerably among occupational groups. The highest percentage of men who had deductions was in the highest income group: professionals, managers, officials, and proprietors.

Health Costs. The average fixed health costs excluding health insurance were \$102.73 when all families were analyzed. Of the 28 percent who had any health costs considered fixed or regular, the average cost was \$315 (Table 1).

Families in different income levels having fixed health costs were as follows:

Under \$2500	15.1 percent
\$2,500 through \$4,999	21.8 percent
\$5,000 through \$7,499	34.0 percent
\$7,500 through \$9,999	30.3 percent
\$10,000 and over	27.5 percent

Fixed health costs, when they did occur, varied significantly by income level but the relationship was not linear. The highest income families, with \$10,000 and over, averaged \$417 in fixed health costs followed by the lowest income families, with under \$2500, who averaged \$404 in fixed health costs. This leads to some interesting speculation as to what is cause and what is effect.

Total Medical Costs. Total fixed medical costs composed of payments for health insurance, payroll deductions for health insurance, and other fixed health costs were analyzed and shown in Table 3. The average for the sample families was \$149. The average costs for all families ranged from \$77 for families with income under \$2500 to \$219 for families with \$10,000 or more income.

TABLE 2

## Payroll Deduction by Husband's Occupational Group

Occupational group	Total number	Number having deduction	Mean for those having deduction <sup>a</sup>	Standard deviation
1. Professionals, managers, officials, and proprietors.	92	32	\$212	\$119
2. Technicians, clerical, and sales workers.	96	25	\$212	\$142
3. Skilled manual laborers.	193	42	\$194	\$101
4. Operatives and semi-skilled laborers.	161	25	\$161	\$126
5. Low level service and private household workers.	108	13	\$155	\$ 73
6. Unskilled laborers.	148	22	\$181	\$101
7. Military	21	1	\$312	--
All families	819	160	\$191	\$114

<sup>a</sup>F ratio = 1.053, not significant beyond the .05 level.

Average medical costs for various income levels were compared. Significant differences were found when all families were included in the analysis but not when just those families having medical costs were included. The over-all mean for these families was \$302. The percentage of families at different income levels who had any costs was as follows:

Under \$2500	21.0 percent
\$2,500 through \$4,999	31.1 percent
\$5,000 through \$7,499	55.6 percent
\$7,500 through \$9,999	56.0 percent
\$10,000 and over	64.5 percent

Among the low-income families, 105 or about 28 percent had some fixed medical costs (Table 3). These families with income below \$2500 averaged \$377 for medical costs and families with income from \$2500 to \$5000 averaged \$305.

#### Influence of Selected Variables Upon Amount of Insurance Payments

The multiple regression step-wise technique was applied to find the association of selected variables with the fixed expenses for health insurance. Fifteen variables were selected because they were available in the data collected. Through stepwise regression, seven variables remained as significant contributors at the .10 level and yielded a multiple R of .31.

The standardized multiple regression equation was  $Y = .33X_1 - .25X_2 + .15X_3 + .12X_4 + .09X_5 + .06X_6 + .04X_7$  in which

Y = expenses for health insurance  
 $X_1$  = income index  
 $X_2$  = total disposable income  
 $X_3$  = husband's age  
 $X_4$  = life insurance  
 $X_5$  = wife's years of schooling  
 $X_6$  = fixed health costs  
 $X_7$  = wife's disability index

TABLE 3

## Total Fixed Medical Costs by Income Level

Income level	Medical costs				Standard deviation
	Total families		Families with costs		
	Number	Mean <sup>a</sup>	Number	Mean <sup>b</sup>	
\$0 through 2,499	118	\$ 77	24	\$377	\$396
\$2,500 through 4,999	257	\$ 96	81	\$305	\$466
\$5,000 through 7,499	357	\$164	199	\$294	\$264
\$7,500 through 9,999	234	\$139	131	\$249	\$213
\$10,000 and above	273	\$219	176	\$340	\$403
Total	1239	\$149	611	\$302	\$338

<sup>a</sup>F ratio 9.011, significant beyond the .01 level.

<sup>b</sup>F ratio 1.701, not significant beyond the .05 level.

In terms of standardized partial regression coefficients (betas), income index (an objective measure of income adequacy) was the most important in explaining expenses for health insurance. As income index increased, families allocated more of their income to life insurance. Total disposable income, which does not take into consideration the cost of food for the given composition of the family, was negatively associated. Variables eliminated from the equation were number in the family, expenses for car insurance, husband's disability index, income dependency, not enough money for doctor or dentist, husband's occupational level, perceptions of income adequacy, and husband's years of schooling.

The finding that both fixed health costs and woman's disability were positively associated with expenses for health insurance has relevance to the proposals for health insurance premiums regardless of the consumer's medical history or risk. An insurance orientation for some families may have been indicated by the finding that life insurance was positively associated with expenses for health insurance ( $r = .14$ ).

Health insurance expenses by families themselves were examined in relation to payroll deductions for health insurance. The Pearson-product moment correlation coefficient ( $r = -.10$ ) was significant beyond the .01 level. As payroll deductions increased, non-payroll expenses for health insurance decreased;



although little of the variance in health insurance expenditures by families themselves could be explained by the payroll deductions.

Possible changes for families who were not paying out of their own finances for health insurance would be increases for all under the Health Security Plan and increases for high income families under the Comprehensive Health Insurance Plan (Figure 1).

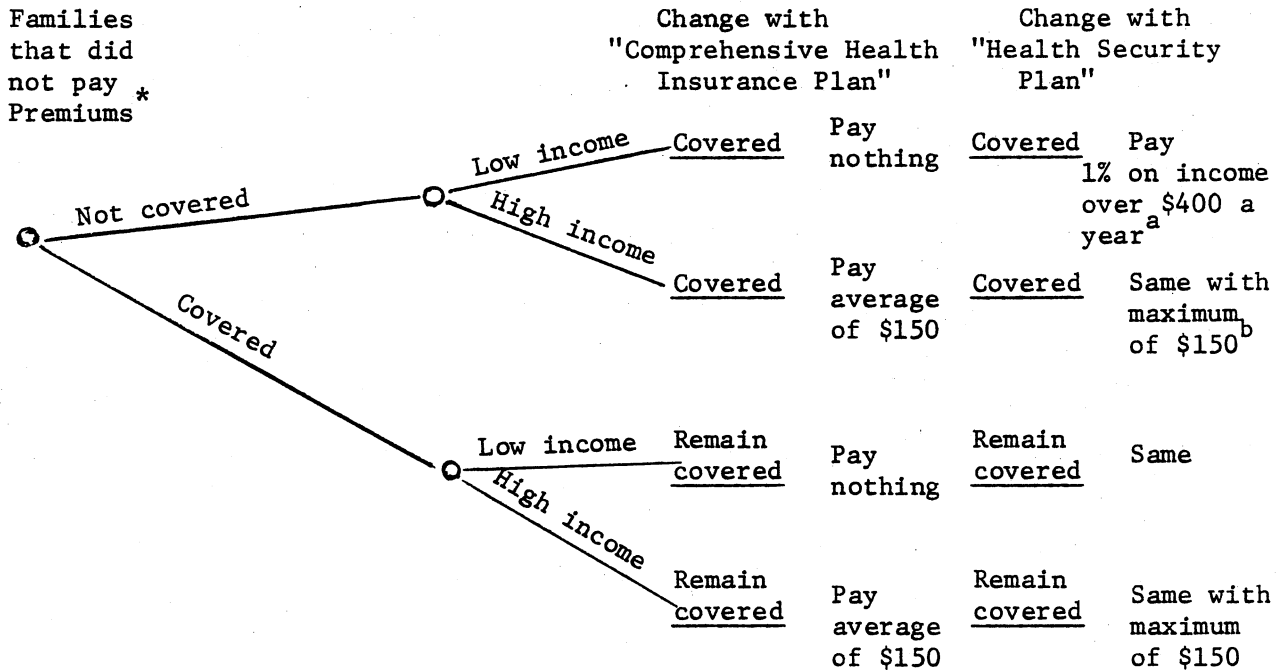


Figure 1. Model of possible changes families would make under proposed health plans.

<sup>a</sup>Excluding certain types of income specifically excluded from the other taxes.

<sup>b</sup>One percent of \$15,000.

Some of the families currently not paying for health insurance are covered by insurance through work-related benefits. Other families are not currently covered but would be covered under a proposal when adopted. If these other families are low-income, below \$5,000, they would continue to not pay under the "Comprehensive Health Insurance Plan" but would pay one percent of \$15,000 under the "Health Security Plan" unless the type of income was specifically excluded from the other taxes (Congressional Record, January 31, 1973, p. 30). Families with high income, regardless of their past coverage, would have to allocate money for insurance premiums that heretofore they had not done. Under the "Comprehensive Health Insurance Plan," their payroll deductions would average \$150 with the employer contributing an average of \$450. Under the "Health Security Plan," one percent of their income in addition to current social security contributions would be paid on 125 percent of the social security wage base.

Comparison of Current Medical Costs Compared to Estimated Costs. Which proposed plan most closely matches what families were paying for fixed medical costs? As shown in Table 4, at every income level, families as a whole were paying more than the cost estimates for the Health Security Plan. The overall average paid was nearly the same as the cost estimate for the EHIP, but families with income under \$7,500 were paying more and families with \$7,500 and over were paying less. Under AHIP families with income under \$5,000 would pay nothing whereas families with \$5,000 or over would pay considerably more than they did pay. The projected average would be \$600 compared to the overall average of \$149 paid by the sample families.

Of families in the lowest income group, the majority (80 percent) did not have any medical expenses considered fixed. Under the Health Security Plan many of these would pay one percent of their income. For those families who had had fixed expenses which averaged \$377 this reduced fixed cost would be extremely helpful. Under EHIP and AHIP these lowest income families would pay nothing. A similar change would occur for families in the \$2,500 through \$4,999 level.

Of families in the \$5,000 and \$7,500 levels, the minority having no fixed expenses would begin paying one percent under the Health Security Plan and \$105 and \$210, respectively, under the EHIP. These amounts would be less than the average of fixed medical expenses (\$294 and \$249) for families having such expenses. Under AHIP they would have to pay more than their average fixed expenses. Of families in the \$7,500 through \$9,999 level, they would pay an average of \$600 compared to an average \$249 for those who had any expenses and none for the others.

Of families with \$10,000 and over income, the 22 percent that heretofore did not have fixed expenses would pay \$100 to \$150 under Health Security Plan, \$315 under EHIP and \$900 under AHIP. For families with fixed expenses averaging \$340 and ranging from \$20 to \$3400 some would pay considerably less under any of the three estimates. Many of those working in high risk occupations or who were health risks would be under AHIP and would pay more than previously.

#### Summary and Conclusions

The current cost of families' fixed payments for medical care was compared to estimated fixed payments under proposed health plans. Families selected randomly from six metropolitan disadvantaged areas were paying an average of \$149 per year for total fixed medical costs. The average for just those families who had any fixed medical costs was \$305. In contrast, estimates for insurance costs under the Health Security Plan averaged \$84 and under the Comprehensive Health Insurance Plan averaged \$150 for those qualifying for the Employee Health Insurance Plan and \$600 for those qualifying for the Assisted Health Insurance Plan.

Compared to cost estimates for the proposed Health Security Plan, families at every income level, on the average, were paying more for fixed medical expenses. Compared to the estimates for EHIP in the proposed Comprehensive Health Insurance Plan, families on the average were paying the same amount. All families under \$7,500 income were paying more than they would under proposed EHIP costs but families over \$7,500 were paying less. However, the families that had fixed medical costs were paying considerably more than estimated costs under EHIP. Compared to AHIP of the Comprehensive Health Insurance Plan, families over \$7,500 income were paying considerably less.

TABLE 4

Current and Estimated Annual Costs (fixed) for Families Below Age 65 by Income Group

Total disposable annual income range	Current Costs			Estimated Costs	
	Average income	Costs of families in sample for insurance and medical costs	Distribution of families No. Percent	Cost estimates for the "Comprehensive Health Insurance Plan" <sup>a</sup> AHIP (\$0)	Cost estimates for the "Health Security Plan" <sup>b</sup> EHIP (\$0)
(\$0 - 2,499)	(118)	(10)	(96)	0	0
					(1 percent on earned and unearned income except for specified types)
	\$ 0	94	80		
	\$ 1 - 149	9	8		
	150 - 299	3	3		
	300 - 599	7	6		
	600 - 899	3	3		
	900 and above	2	2		
(\$2,500 - 4,999)	(257)	(21)	(96)	0	25 - 49.99
	177	68			
	0	30	11		
	1 - 149	24	9		
	150 - 299	17	6		
	300 - 599	7	2		
	600 - 899	2	1		
	900 and above	2	1		
(\$5,000 - 7,499)	(357)	(29)	(164)	300	105
	158	44			(35% by employee 65% by employer)
	0	70	19		
	1 - 149	56	15		
	150 - 299	47	13		
	300 - 599	21	5		
	600 - 899	5	1		
	900 and above	5	1		
(\$7,500 - 9,999)	(234)	(19)	(139)	600	210
	103	44			(35/65)
	0	51	21		
	1 - 149	39	16		
	150 - 299	34	14		
	300 - 599	4	1		
	600 - 899	3	1		
	900 and above	3	1		
(\$10,000 and above)	(273)	(22)	(219)	900	315
	97	35			(35/65)
	0	54	19		
	1 - 149	52	19		
	150 - 299	46	16		
	300 - 599	17	6		
	600 - 899	7	2		
	900 and above	7	2		
(All families)	\$8435	(1239)	(149)	600	150
		100)			84.35

a The White House "Fact Sheet. The Comprehensive Health Insurance Plan," Office of the White House Press Secretary, February 6, 1974, p. 6.  
 b Congressional Record, 93d Congress, First Session, Vol. 119, No. 17, January 31, 1973, p. 30.

In this sample, 51 percent of the families did not consider any expenses, fixed or regular, for medical costs. Forty-three percent of these families were low income and, therefore, would not pay under proposed AHIP and EHIP but all would pay under the Health Security Plan unless income was specifically excluded as a determinant.

Health insurance, other than payroll expenses, was paid by 17 percent of the families themselves. The average spent on health insurance premiums was \$212 for those that did pay. Variables associated with amount spent on health insurance, in order of importance, were income index, total disposable income (negatively), husband's age, life insurance, wife's years of schooling, fixed health costs, and wife's disability index but not number in family, expenses for car insurance, husband's disability, income dependability, perceptions of income adequacy and husband's years of schooling. Income index took into account the composition and number in the family as well as total disposable income.

Payroll deductions for health insurance were not commonly found among families. Fifteen percent had such deductions. Under the EHIP the employee would deduct 35 percent of the insurance payment and the employer would contribute 65 percent. Heretofore, for these sample families most employers contributed all or none. Adjustments would be made in this realm, especially for the 85 percent of the families not having had deductions. For those having deductions, the \$185 average is higher than the proposed \$150.

Fixed or regular health costs other than health insurance occurred for 28 percent of the families. Their average cost was \$315.

Total fixed medical costs which were composed of expenses for health insurance, health costs, and payroll deductions that were fixed or regular for families varied with income levels. Families below \$2,500 income averaged total costs of \$77; those between \$2,500 and \$5,000 averaged \$96; those between \$5,000 and \$7,499 averaged \$164; those between \$7,500 and \$9,999 averaged \$139; and those with \$10,000 and above averaged \$273. The low averages occurred because only about half of the families had such expenses.

The immediate impact upon families as consumers will probably be the change in their budgets for fixed allocations for medical expenses. The increased benefits or coverage will be realized as programs are implemented and utilized by families. Although the sample families at every income level, on the average, were paying more for fixed medical expenses than the estimates for the proposed Health Security Plan, over half had no expenses as fixed. Those families would lose their flexibility in health expense allocations under this plan. Many of the other families would pay considerably less.

Families under \$5,000 income would pay nothing for the Comprehensive Health Plan which would be considerably less for 44 percent of the families and no change for the others. Families with \$5,000 to \$7,500 incomes, on the average, would pay about the same under AHIP but less under EHIP than they were paying for fixed health care. Of course the 29 percent not having fixed expenses in this income group would have considerable increases in fixed expenses. The greatest impact upon budget allocations would be for families with incomes over \$7,500 who would qualify for AHIP. These cost estimates are much higher than they paid previously.

## FOOTNOTES

1. This study contributes to inter-regional Agricultural Experiment Station research project NC-90, "Factors Affecting Patterns of Living of Dis-advantaged Families." Cooperating states: California, Hawaii, Illinois, Indiana, Iowa, Kansas, Missouri, Nebraska, Nevada, Ohio, Texas, Vermont, Wisconsin. Journal Paper No. 5643, Purdue Agricultural Experiment Station, West Lafayette, Indiana 47907.
2. Flora Williams is Assistant Professor, Home Management and Family Economics Department, School of Home Economics, Purdue University.

## REFERENCES

- Congressional Record. 93d Congress, First Session, January 31, 1973, 119(17), 30.
- Congressional Record. 93d Congress, Second Session, February 6, 1974, 120 (12), S1341.
- Orshansky, M. How poverty is measured. Monthly Labor Review, February, 1969, 92 (6), 32-62.
- Reiss, A. G. Occupation and social status. New York: Free Press of Glencoe, 1961.
- United States Bureau of the Census. Maps of major concentrations of poverty in standard metropolitan statistical areas of 250,000 or more populations. Vol. 1. Prepared for the Office of Economic Opportunity. Washington, D.C.: U.S. Government Printing Office, 1966, iv.
- United States Bureau of the Census. Statistical abstract of the United States: 1971. (92d edition) Washington, D.C., 1971.
- The White House. Fact sheet: The comprehensive health insurance plan. Office of the White House Press Secretary, February 6, 1974, 6.

TELEVISION ADVERTISING: A STUDY ON ITS APPLICABILITY  
FOR PROMOTING THE DISADVANTAGED TO EMPLOYERS

James E. Haefner  
University of Illinois at Urbana-Champaign

A field study was made in three cities of Illinois: Peoria, Rockford and Decatur. Ten television spot announcements dealing with the hiring and retraining of disadvantaged individuals were shown during prime television time for six weeks in Peoria and Rockford. The spots were not shown in Decatur. Pre-measurements of the behavioral intentions of employers to hire or retrain disadvantaged people were taken simultaneously in all three cities. Post-measurements of intentions were gathered in all three cities after the ad campaign was completed. In addition, measures of message recall and comprehension were collected in the Peoria and Rockford areas. Recall and comprehension rates for the spot announcements were very high for employers in Peoria and Rockford, much above many past informational advertising campaigns. Negative shifts in behavioral intention measures occurred, however. The most likely explanation for these results can be found in reactance theory. When an individual perceives that his alternatives are being restricted, he will tend to view the object or stimulus that is attempting to restrict his alternatives in a more negative manner. The spot announcements appeared to be viewed by employers as an attempt to restrict alternatives in the area of hiring and retraining. The study indicates the need for less "hardsell" messages in this area of informational advertising.

#### Introduction

This paper presents the results of a field study made in three cities of Illinois: Rockford, Peoria and Decatur. A six-week, informational, television-advertising campaign was conducted to improve employer knowledge and behavioral intentions toward the hiring and training of groups of disadvantaged adults. The advertisements were shown in Rockford and Peoria while Decatur was the control community. Interviewing before and after the campaign was conducted in each of the three cities to determine what effects, if any, the campaign had on employers.

#### Effectiveness of Past Informational Advertising Campaigns

Studies of past informational campaigns reveal that: (1) relatively few informational campaigns have been evaluated, and (2) when systematic tests of the effectiveness of informational campaigns have been run, the results have been discouraging (Haskins, 1968; Klapper, 1960; Douglas, 1970; Salado and Scherer, 1973). For example, traffic safety campaigns, such as "Buckle Up for Safety," and "Watch Out for the Other Guy," were judged to be inconclusive, although over \$40 million in advertising space and an unknown amount of creative and

administrative effort were expended (Salado and Scherer, 1973). In another campaign, ads were developed to inform the audience about the United Nations and its functions. The campaign produced no increase in audience knowledge about the United Nations (Star, 1950).

This is not to say, of course, that all informational advertising campaigns have been unsuccessful. Some examples of the few successful campaigns include Kate Smith's selling of more than 39 million dollars of bonds in 18 consecutive hours (Kotler, 1969); Douglas and others' (Douglas, 1970) ability to change community attitudes toward mental retardation; and Salcedo and others' (Salcedo and Reed, 1974) ability to increase audience knowledge of, and strengthen attitudes toward, pesticide safety and the pesticide label.

Overall, however, informational advertising campaigns have generally been unsuccessful in achieving their given goals: developing of awareness, imparting of information, changing of attitudes, or bringing about action. Why the great failure in informational advertising while more traditional product-oriented advertising has been more successful? The reasons are several. First, most, if not all, of these campaigns have relied entirely upon public service time and space to disseminate their messages. These public service announcements are placed on the air at the discretion of the public service director or station manager, and more likely than not they are aired at 6:30 a.m. or at 11:30 p.m. It is highly unlikely that any significant number of the intended audience will be watching the station at these times. Gruenhagen (1969) in a recently completed study indicated that one of the reasons his campaign failed was the lack of prime time exposure his ads were given.

A second reason for the failure of these informational campaigns is that they have often dealt with firmly entrenched attitudes and values. Attempting to instill new attitudes or values in such areas such as sex education, racial relations, fluoridation, or birth control present a much more difficult problem for the communicator than attempting to reshape existing attitudes. For example, the seller of toothpaste does not have to socialize persons into new dental care habits, but rather into which brand of a familiar and desired product to purchase (Kotler, 1971, p. 6). This is not meant to imply that instilling new attitudes or behavioral patterns is impossible, but that it requires a more sophisticated and concentrated effort than many traditional product promotions.

A third reason for the poor results of informational campaigns is failure on the part of the communicator to recognize that he may have to do more than just advertise his idea or service. The great success of the rightist, Father Coughlin, in the thirties was due to a combination of an ad campaign (Coughlin's addresses on nationwide radio), plus the coordinated distribution of newspapers, pamphlets, and locally organized face-to-face discussions among relatively small groups (Lazarfeld, 1949). What is needed is a sound marketing strategy and not just a sound advertising strategy.

In summary, the review of the literature reveals the following:

1. Few informational campaigns have been evaluated.
2. When informational campaigns were evaluated, the results were often negative.
3. Failure of these campaigns has probably resulted from their sole reliance

on public service spots, their attempt to deal with issues that require social reconditioning, and their reliance on advertising alone to create the desired changes.

## Method

### Subjects

Employers were operationally defined to include respondents who could hire or were in a position to recommend the hiring of an individual. As a result of this definition of employers, potential respondents were divided into three groups: 1) president or owner, 2) personnel manager or personnel director, and 3) foreman or supervisor. The sample for each city contained an adequate number of individuals for each of three sub-classifications.

A total of 105 (Peoria), 101 (Rockford), and 80 (Decatur) employers were contacted for the pre-measurement phase of the project. Sample sizes for the post-measurements were Peoria (94), Rockford (93), and Decatur (74).

### Sampling

For each city, a directory of area employers was obtained. Those firms employing fewer than five people were excluded. From the remaining list, 100 firms were selected in Peoria, Rockford, and Decatur with probabilities proportioned to size (measured by number of employees). For each selected firm, the number of interviews to be completed with management personnel was assigned proportionally to the size of the firm. The total sample for each city was then randomly divided into three equal groups designated high level (president, owner); medium level (personnel manager, personnel director); low level (foreman, supervisor).

All employers in the final sample were contacted by telephone for their pre-measurement responses. The same group of employers was again contacted by telephone for their post-measurement responses.

### Stimulus Material

Ten color television spot announcements, five were 60 seconds in length and five were 30 seconds in length, were developed. Two major themes evolved: one theme indicated the importance of employers hiring the disadvantaged; and the second theme dealt with the need to retrain and upgrade the existing skills of disadvantaged people already employed by them. Disadvantaged adults were portrayed in the ads as those who might be discriminated against because of age, sex, or race. At the end of each message, the seal of the State of Illinois appeared.

### Design

A before-after with control group was utilized as the experimental design. Peoria and Rockford were designated as the test cities while Decatur was the control community.

### Pre-Measurement

Each employer was asked a series of demographic items dealing with such areas



as age, education, length of employment, position within company, and present training programs for employees. In addition, behavioral intention questions modeled after Fishbein (1972) were also asked.

#### Post-Measurement

The post-measurement questionnaire for employers was the same as the pre-measurement questionnaire with three changes: 1) the demographics collected earlier were not collected again, 2) a section was added to measure advertising recall, and 3) a section was added to measure message comprehension.

The recall questions asked respondents whether they remembered seeing any television ads that had the themes, "When They Succeed, We All Do;" and "Take Another Look." If a respondent answered in the affirmative, he was then asked what he thought the messages were trying to say (message comprehension). For those respondents who indicated that they were not sure whether they saw the messages, they were given a series of statements which described each of the ads designed in such a way that the description would have not indicated what the themes of the messages were. The list also contained some statements describing fictitious ads and was thus used as a checking device to prevent respondents from saying they had seen all the ads. If a respondent stated he remembered seeing the fictitious ads, he was not counted as having recalled the ads.

#### Procedure

Departing from the usual practice of informational advertising campaigns, it was decided not to rely on free public service spots for the ad campaign. Instead, time was purchased at the local affiliates of ABC, CBS, and NBC in Rockford and Peoria (the experimental cities). Table 1 reports on the frequency schedule of the ten announcements.

TABLE 1

Frequency Schedule of Disadvantaged Television Announcements

Schedule Time	Number Per Week	Length of Commercial
7 a.m. to 9 a.m.	3	60 seconds
3 p.m. to 5 p.m.	3	60 seconds
Evening news	3	60 seconds
Evening prime time	6	60 seconds
Night news	3	60 seconds
Late night	3	60 seconds
Saturday afternoon sport	1	30 seconds
Saturday night prime time	1	30 seconds
Sunday afternoon sport	1	30 seconds

Note. The frequency schedule was the same for both Rockford and Peoria. The ads were shown on the local affiliates of ABC, CBS, and NBC, and were run from September 3, 1973 to October 14, 1973.

Simultaneous pre-measures were taken before the campaign began in all three cities and simultaneous post-measures were taken in all three cities directly after the campaign.

## Results

### Sample Statistics

Table 2 presents some summary sample statistics for employers in each of the three cities.

TABLE 2

Summary Statistics of First Wave Employer Sample: Rockford, Peoria and Decatur

Summary Statistic	Rockford		Peoria		Decatur	
	Frequency or mean	Pct.	Frequency or mean	Pct.	Frequency or mean	Pct.
1. Completed Interviews	101	-	105	-	80	-
2. Managerial Level in Firm:						
a. Top Management (President, Vice President, etc.)	25	24.8%	36	34.3%	32	40.0%
b. Middle Management	60	59.4%	51	48.6%	35	43.8%
c. Lower Management (First line foreman, etc.)	16	15.8%	18	17.1%	13	16.2%
3. Sex						
a. Male	80	79.2%	92	87.6%	73	91.2%
b. Female	21	20.8%	13	12.4%	7	8.7%
4. Average Age	41		46		46	
5. Could Hire Personnel						
a. Yes	86	81.9%	83	82.2%	73	91.2%
b. No	19	18.1%	18	17.8%	7	8.7%
6. Average Educational Level (where 1=Grammar School, 2=some High School, 3=High School Grad, 4=some College, 5=College Degree 6=Post Grad Degree)	4.1	-	3.9	-	4.0	-
7. Average Number of Years Worked for Company	13	-	17.8		16.6	-

Adequate numbers of top, middle, and lower management personnel were obtained from each city. The employer respondents were with the firm in question for a fairly large number of years, thus indicating that they understood the hiring and training mechanisms of the company. It is also interesting to note the fairly large number of women respondents that were obtained. Other demographic data collected along with those presented in the paper indicated that the samples were homogenous enough to permit comparisons amongst them. No data is presented for the second wave sample as they were very similar to the first wave respondents.

Message Recall and Comprehension Frequencies for Peoria and Rockford

Table 3 presents the results of recall and comprehension for the campaign.

TABLE 3

Recall and Comprehension Frequencies for Peoria and Rockford  
(Second Wave)

Samples	Peoria			Rockford			Total		
	Recall		Comprehension	Recall		Comprehension	Recall		Comprehension
	Frequency %	Frequency %	Frequency %	Frequency %	Frequency %	Frequency %	Frequency %	Frequency %	
Employers (n=187, Peoria n=94 Rockford n=93)	69	<u>73.3</u>	1 43 45.7	63	<u>67.7</u>	1 35 37.6	132	<u>70.6</u>	1 78 41.7
			2 4 4.3			2 8 8.6			2 12 6.4
			3 4 4.3			3 2 2.2			3 6 3.2
			51 <u>54.3</u>			45 <u>48.4</u>			96 <u>51.3</u>

- Note. 1 - refers to mentioning hiring theme  
 2 - refers to mentioning training theme  
 3 - refers to mentioning both the hiring and training theme

Recall rates for employers were very strong with 70.6% of the combined Peoria and Rockford samples indicating that they had seen the messages. Although comprehension rates for employers fell to 51.3% for the combined samples, this is certainly a more than acceptable rate for a six week campaign. It should also be noted that the hiring theme was comprehended most clearly for employers; only a small percentage perceived a training theme.

Behavioral Intention Results

Table 4 presents the results of the test conducted on two behavioral intention questions (willingness to hire and willingness to retrain the disadvantaged). Only those employers who both recalled and comprehended the messages were included in the analysis.

TABLE 4

Mean Change Score Tests for Behavioral Intention Questions Administered to Employers; (Peoria, Rockford, and Decatur)

Cities	Mean Change Score	Variance	t Value	Significant at:
<u>Peoria (n=51)</u>				
1. Willingness to hire	-.224	.594	--	--
2. Willingness to train	.021	.717	--	--
<u>Rockford (n=45)</u>				
1. Willingness to hire	-.089	.174	--	--
2. Willingness to train	-.222	.540	--	--
<u>Decatur (n=74)</u>				
1. Willingness to hire	.081	.651	--	--
2. Willingness to train	-.041	.505	--	--
Peoria vs. Decatur (willingness to hire)	--	--	-15.912	.01
Peoria vs. Decatur (willingness to train)	--	--	3.040	.01
Rockford vs. Decatur (willingness to hire)	--	--	-11.883	.01
Rockford vs. Decatur (willingness to train)	--	--	- 9.549	.01

Note. Since employers were administered two behavioral intention questions, one dealing with willingness to hire and the other dealing with willingness to train, two sets of mean change scores are reported.

Significant changes in intentions to hire and retrain occurred in all cases. In the Peoria vs. Decatur (willingness to hire), Rockford vs. Decatur (willingness to hire), and Rockford vs. Decatur (willingness to retrain), all significant changes were negative; behavioral intention scores decreased from pre to post-measurement. Only for Peoria vs. Decatur (willingness to retrain) was there a significant positive change in intention scores.

#### Discussion and Implications

The strong message recall and comprehension rates for the campaign were a result of the purchase of television time rather than relying upon public service information spots. With the purchase of time, it was possible to ensure that the target market of employers would be adequately reached.

Only one positive shift in behavioral intentions occurred (Peoria vs. Decatur - willingness to retrain). This shift had limited implications as a few members of the Peoria employer sample indicated very large shifts while the greater part of this sample indicated no shifts or small negative shifts.

The more interesting finding dealt with the constant negative shifts that

occurred in intentions for all other measurement situations. An explanation for this behavior can be found in Brehm's work (1966). Psychological reactance theory was formulated to explain why some people respond negatively to any force which they perceive restricts their freedom of action:

"...when a person believes himself free to engage in a given behavior, he will experience psychological reactance if that freedom is eliminated or threatened with elimination. Psychological reactance is defined as the motivational state directed toward the re-establishment of the threatened or eliminated freedom, and should manifest itself in increased desire to engage in the relevant behavior...."

Reactance theory research has produced unequivocal results in the literature (Biondo, 1971; Grebitz-Gniech, 1971; Wicklund, 1968) as well as indicating important marketing applications (Mazis, 1973; Venkatesan, 1966).

It may be that the ads produced a reactance or "boomerang effect" on employers, because they perceived the messages as a "hardsell" that attempted to threaten their freedom of choice in the hiring and retraining areas. Three reasons can be given in defense of the above reasoning: (1) Interviewers on the project indicated that many of the employer respondents stated that they were tired of being told what they should do with disadvantaged members of their community. (2) If an individual perceives that a stimulus is threatening a reduction in choice alternatives, reactance effects can be produced (Brehm, 1966, pp. 71-90). Thus actual removal of choice alternatives need not occur to produce psychological reactance. (3) Since the seal of the State of Illinois appeared after each ad, employers may have perceived the sender of the message to have more legitimate power than the sender actually did. The message sender, therefore, may have been perceived to have the power to follow up on the message with more tangible actions.

Two major alternatives can be offered to the "hardsell" communication tactic. First, messages should be developed which mainly discuss the benefits to the company of hiring or retraining disadvantaged people. This "softsell" approach could be combined with other information in an ad that would indicate that some agency (local, state, or federal) would be available to assist in the initial implementation of a disadvantaged hiring or retraining program. With the message and source being perceived more in a coordinating role rather than in an enforcement role, psychological reactance will be less likely to occur.

Secondly, it may be that the advertising function alone will not be able to bring about significant behavioral changes in this area. A social marketing program (Kotler, 1971), with advertising but one of its tools, may provide more positive results. Recognition of the costs involved in hiring the disadvantaged, the problems in obtaining disadvantaged community members when an employer needs them, as well as the need to communicate effectively with employers may provide more effective results in this area of societal concern.

#### FOOTNOTES

1. This researcher is indebted to the Division of Vocational and Technical Education, State of Illinois, for providing funds to conduct this research. Only a small part of the results of this study will be presented.
2. James E. Haefner is an Assistant Professor of Advertising at the University of Illinois, Urbana-Champaign (103 Gregory Hall).

## REFERENCES

- Biondo, J. & MacDonald, A. P. Internal-external locus of control and response to influence attempts. Journal of Personality, 1971, 39, 407-19.
- Brehm, J. A theory of psychological reactance. New York: Academic Press, 1966, 15-16.
- Douglas, O., Westley, B., & Chaffee, S. An informational campaign that changed community attitudes. Journalism Quarterly, 1970, 47, 479-487, 492.
- Fishbein, M., & Ajzen, I. Attitudes and opinions. In P. Massen and M. Rosenzweig (Eds.), Annual Review of Psychology, Vol. 23. Palo Alto, California: Annual Reviewing Inc., 1972, 487-544.
- Greibitz-Gniech, G. Some restrictive conditions for the occurrence of psychological reactance. Journal of Personality and Social Psychology, 1971, 19, 188-195.
- Gruenhagen, R. The effects of planned communication program on change of attitudes and knowledge of the urban dweller toward chemicals and pesticides. Virginia Cooperative Extension Service, Virginia Polytechnic Institute, November, 1969.
- Haskins, J. Evaluation of effects of educational and mass communication techniques. 1968 National Safety Congress Transactions, 24, 77-109.
- Holmes, J. G., & Strickland, L. Choice freedom and confirmation of incentive expectancy as determinants of attitude change. Journal of Personality and Social Psychology, 1970, 14, 39-45.
- Klapper, J. The effects of mass communication. New York: The Free Press, 1960.
- Kotler, P., & Zaltman, G. Social marketing: An approach to planned social change. Northwestern University. Unpublished working paper, 1969.
- Kotler, P., & Zaltman, G. Social marketing: An approach to planned social change. Journal of Marketing, 1971, 35, 3-12.
- Lazarfeld, P. & Merton, R. Mass communication, popular taste, and organized social action. In W. Schramm (Ed.), Mass communication. Urbana, Illinois: University of Illinois Press, 1949, 459-480.
- Mazis, M., Settle, R., & Leslie, D. Elimination of phosphate detergents and psychological reactance. Journal of Marketing Research, 1973, 10, 390-395.
- Salcedo, R., Reed, H., Evans, J., & Kong, A. A successful information campaign on pesticides. Journalism Quarterly, 1974, 51, 91-95.
- Salcedo, R., Scherer, C. W., & Alison, K. Effects of a public service informational campaign on audience knowledge of and attitudes toward certain food buying practices and nutrition information. Office of Agricultural Communication, University of Illinois. Working paper #2, 1973, 7.
- Star, S. & Hughes, H. A report on an educational campaign: The Cincinnati plan for the United Nations. American Journal of Sociology, 1950, 55, 389-400.

Wicklund, R. & Brehm, J. Attitude change as a threat to attitudinal freedom.  
Journal of Experimental Social Psychology, 1968, 4, 64-75.

Venkatesan, M. Consumer behavior: conformity and independence.  
Journal of Marketing Research, 1966, 3, 384-387.

SOCIAL SCIENCE INPUTS TO PUBLIC POLICY FORMATION:  
MASSACHUSETTS AND THE UNIT PRICING REGULATIONS

Saul Barry Wax<sup>1</sup>  
Massachusetts Consumers' Council

The Consumers' Council regulates unit pricing on a state-wide basis. The Council constantly explores ways in which unit pricing information may be made both more accessible and more meaningful to consumers. In June, 1974, the Council held public hearings, and based upon the testimony presented at the hearing promulgated amended unit pricing regulations. A fundamental alteration in the regulations was an increase in the size of the specified units of measure/or regulated commodities. The change was based upon testimony describing the applicability and accuracy of the Weber-Fechtner law to consumer utilization of unit pricing. The research conducted and data presented to the Council by a social psychologist represent the clear use of social science data in a public policy making arena.

The Consumers' Council is an independent consumer advocacy agency founded by the General Court (legislature) in 1963. (See Appendix A.) The Council consists of thirteen members, eight of whom are appointed by the governor (public members) for terms coterminus with his/hers. The remaining five members are ex-officio: The Attorney General, Department of Banks and Banking, Department of Public Utilities, Department of Labor and Industries and the Department of Insurance. All members of the Council have one vote and a Chairman is appointed by the Governor.

The Council acts as an ombudsman within the governmental structure: Its primary functions are consumer advocacy before both the General Court and regulatory agencies, consumer education, and the drafting and filing of pro-consumer legislation. The Council is a close approximation of what the Agency for Consumer Advocacy is proposed to be at the federal level.

Although the Council is fundamentally an advocacy agency, in 1970 the General Court, in an attempt to resolve a prickly issue, delegated regulatory authority over unit pricing to the Consumers' Council. (See Appendix B.) Pursuant to the provisions of the statute, the Council held public hearings and adopted a set of regulations. Several years and several sets of regulations later, the Council amended the regulations yet again by specifying acceptable units of measure for commodities which have to be unit priced.

The Problem: Impact and Utility of a Public Policy

The Council promulgated unit pricing regulations with the presumption that such information would enable consumers to save money on their grocery bills. No doubt, the unit price of items does provide valuable comparative cost information. However, the availability of useful data does not insure usage. The proverb about leading a horse to water is surely an accurate one.



The supermarket is a veritable carnival row of sights, sounds and colors, and, consequently, shopping challenges the consumer to separate the hype from those things which help. The Massachusetts unit pricing label was designed to maximize clarity and readability; the size of the type, and the placement of the information and the color of the label were selected with a mind toward distinguishing the unit price labels from the deluge of self-serving information found in the supermarket. Studies had been conducted in a number of areas where several chains had voluntarily implemented unit pricing (Friedman, 1971, 1972; Consumer Research Institute, 1971). A distillation of the findings revealed several bits of information. First, usage of unit pricing hovered at around twenty percent. For the most part, those who used unit pricing were disproportionately white, better educated and middle class.<sup>2</sup> In many instances, unit price labels were difficult to read, confusing, illegible or simply wrong. And these latter factors took their toll of potential users. To be sure, within a supermarket, there is nothing more discouraging than to try and make use of a poorly displayed label.

The Consumers' Council had by June, 1974, progressed as far as it could in the domains of label design, color specification, type size, authority over means of display, and the positioning of information. If unit pricing information was to be improved, the change would have to be made in an area which heretofore the Council had not taken into consideration.

#### Unit Pricing and The Weber-Fechtner Law

The Weber-Fechtner Law holds that as the numerical difference between figures increases so does the psychological impact of that difference. The difference between 8 cents per ounce and 10 cents per ounce is not appreciable, but the difference between \$1.28 and \$1.60 per pound is quite noticeable. In writing the next amended version of the unit pricing regulation, the Council turned to research which investigated the impact of the Weber-Fechtner Law.

In May, 1974, a wire story article appeared in the pages of the Boston Globe. The gist of the story, as edited by the Globe personnel, was that unit pricing was simply not working. This bit of information was contrary to the Council's data and hence prompted an investigation into the research cited in the article. A quick check revealed that the story had been filed by a reporter working for the Philadelphia Inquirer and was based upon a study conducted by Professors Lloyd Sloan and David Walton (1974) for presentation at the Eastern Psychological Association Convention. Their study and concomitant finding were refreshingly straightforward. An increase in "the unit size used in presenting unit price information significantly affected consumer buying behavior [p. 3]." Upon receipt of the Sloan-Walton study, the authors were contacted and invited to make a presentation at the Council's upcoming public hearing.

The invitation was accepted by Professor Walton and on June 6, 1974, he came to Boston and outlined in appreciable detail their research findings and the apparent implication for public policy formation. Based upon the Sloan-Walton data (which, it should be noted, is consistent with common sense impressions), the Council altered all units of measure for purposes of maximizing the unit price differentials between products. (See Appendix C.)

The Council's action reflects at the very least the responsiveness and willingness of an agency engaged in public policy making to utilize social (behavioral) science data. Social scientists involved in research topics/programs which conceivably could have a bearing upon public policies could profitably note the sequency of events which transpired in Massachusetts.

### Postscript and Projections

Although altering the units of measure for unit pricing may appear to be a minute consideration, the new policy had a significant impact upon a major industry and has reverberated across four states. The changes which were promulgated meant that the provisions of the Massachusetts regulations were inconsistent with the regulations of other states. Consequently, stores which conduct business in several jurisdictions have been and are with the Council's aid petitioning the regulatory authorities in other areas to amend their regulations taking into consideration the recent changes made in Massachusetts.

At the bottom line of this and all other public policies is (or should be) the question of impact and utilization. An assessment of the amended regulations cannot yet be made. The new units of measure necessitated significant changes in stores' computer systems, and most chains are still getting the bugs out.

The continual increases in the price of food make it very likely that consumers will utilize unit pricing more now than ever before. The changes made by the Council underscore the likelihood that such utilization will be easier and more meaningful for consumers.

### FOOTNOTES

1. The author is the Senior Policy Analyst at the Massachusetts Consumers' Council, a State advocacy agency located in Boston, Massachusetts.
2. See especially M. Friedman, 1971.

### REFERENCES

- Consumer Research Institute, Cost of unit pricing in grocery stores. Prepared by the A. T. Kearney and Company, Inc., July, 1971.
- Friedman, M.P. Consumer price comparisons of retail products: The role of packaging and pricing practices and the implications for consumer legislation. Journal of Applied Psychology, 1972, 56, 439-446.
- Friedman, M.P. Dual price labels: Usage patterns and potential benefits for shoppers in innercity and suburban supermarkets. Ypsilanti, Mich.: Eastern Michigan University Center for the Study of Contemporary Issues, 1971.
- Walton, D. & Sloan, L. Increasing the effectiveness of unit pricing: Increasing the unit of unit price. Paper presented at the meeting of the Eastern Psychological Association Convention, Philadelphia, Pa., April, 1974.

APPENDIX A

696

ACTS, 1963. — CHAP. 773.

**Chap. 773.** AN ACT ESTABLISHING A CONSUMERS' COUNCIL.*Be it enacted, etc., as follows:*

SECTION 1. Section 17 of chapter 6 of the General Laws is hereby amended by inserting after the words "control commission", in line 15, as appearing in section 1 of chapter 623 of the acts of 1958, the words: —, the consumers' council.

SECTION 2. Said chapter 6 is hereby further amended by adding after section 114 the following section, under the caption CONSUMERS' COUNCIL: —

*Section 115.* There shall be a consumers' council, hereinafter called the council, consisting of eight members to be appointed by the governor, with the advice and consent of the executive council, for terms concurrent with that of the governor, one of whom shall be a member of the Massachusetts State Labor Council, AFL-CIO and no more than five of whom shall be members of the same major political party, and ex officio, the attorney general, the chairman of the public utilities commission, the commissioner of banks, the commissioner of insurance and the commissioner of labor and industries. The chairman of the council shall be designated by the governor from time to time. Said members shall serve without compensation, but shall receive their expenses necessarily incurred in the performance of their duties. The council shall meet monthly and at other times at the call of the chairman.

The council shall conduct studies, investigations and research and advise the executive and legislative branches in matters affecting consumer interests, co-ordinate consumers' services carried on by departments and agencies, further consumer education, inform the public, through appearances before federal and state committee, commission or department hearings, or otherwise, of such policies, decisions or legislation as are beneficial or detrimental to consumers, inform the governor and the attorney general and other law enforcement agencies of such violations of laws or regulations affecting consumers as its investigations or studies may reveal, and study and report all matters referred to it by the general court or the governor. The council may appear, through its chairman or a member or person designated by him, or through the attorney general, for and in behalf of the people of the commonwealth before boards, commissioners, commissions, departments or agencies of the commonwealth in any hearing or matter affecting the rights of the consuming public or in any proceeding seeking the curtailment of railroad services or an increase of rates or costs of services or commodities, and shall be deemed an aggrieved party for the purpose of judicial or administrative review of any decision or ruling in any such proceedings in which it has so appeared, any other provision of law to the contrary notwithstanding. The council may appoint an executive secretary who shall not be subject to chapter thirty-one, and shall with the approval of the governor and executive council fix his salary. The council may appoint such other assistants, consultants, investigators and experts as it deems necessary to carry out the provisions of this section. Said council shall be provided with offices in the state house or elsewhere in the city of Boston. It may call upon any department, board, commission or officer of the commonwealth or of any political subdivision of the

## ACTS, 1963. — CHAP. 773.

697

commonwealth for such information as it may desire in the course of its duties.

The council may hold public hearings and shall establish rules of procedure governing the conduct of its hearings which shall be made available in printed form to each witness prior to his testimony. Witnesses shall have the right to be represented by counsel and shall before testifying be sworn. The governor may request of the attorney general such legal assistance as may be necessary in carrying out the duties and functions of the council.

*Approved October 31, 1963.*

APPENDIX B

828            ACTS, 1970. — CHAPS. 885, 886.

**Chap. 885.** AN ACT ESTABLISHING A UNIT PRICING LAW FOR CERTAIN RETAIL STORES.

*Be it enacted, etc., as follows:*

SECTION 1. Chapter 6 of the General Laws is hereby amended by inserting after section 115 the following section: —

*Section 115A.* The council may adopt regulations establishing lists of packaged commodities necessary for personal, family or household use to be offered for sale at retail and which may not be sold in retail stores unless there is posted in a conspicuous place at or near the point of sale the price per pound, pint or other unit or measurement of contents and the total sales price. Such regulations shall exempt any packaged commodity whose net weight is one whole unit or two whole units and which has the retail price plainly marked thereon. No packaged commodity shall be included in these regulations which must be individually marked with the cost per unit of weight, liquid, or dry measure, as provided in section one hundred and eighty-one of chapter ninety-four. Said council may adopt such further regulations as are necessary to carry out the intent of this section, provided that a public hearing shall be held relative to any packaged commodity proposed to be regulated. The director of standards shall enforce any regulation adopted pursuant to the authorization contained in this section. Whoever violates any provision of this section shall for the first offence be punished by a fine of not less than ten nor more than fifty dollars, and for a subsequent offence by a fine of not less than twenty-five nor more than one hundred dollars.

Said council shall annually report to the general court on or before the last Wednesday in January of each year relative to any action taken by it pursuant to this section in the preceding year.

The provisions of this section shall not apply to any retail establishment operated by a person as his sole place of business.

SECTION 2. This act shall take effect on January the first, nineteen hundred and seventy-one.

*Approved September 1, 1970.*

---

**Chap. 886.** AN ACT PROVIDING THAT CERTAIN POLICE OFFICERS SHALL BE EXCUSED FROM DUTY WITHOUT LOSS OF PAY WHILE IN ATTENDANCE AS OFFICIAL DELEGATES AT AN ANNUAL CONVENTION OF THE MASSACHUSETTS POLICE ASSOCIATION.

*Be it enacted, etc., as follows:*

SECTION 1. Section 17 of chapter 147 of the General Laws is hereby amended by striking out the last two sentences.

SECTION 2. Section 17D of said chapter 147, inserted by section 2 of chapter 246 of the acts of 1961, is hereby amended by striking out the last sentence.

*Approved September 1, 1970.*

THE COMMONWEALTH OF MASSACHUSETTS,  
EXECUTIVE DEPARTMENT, STATE HOUSE,  
BOSTON, September 1, 1970.

The Honorable JOHN F. X. DAVOREN, *Secretary of the Commonwealth, State House, Boston, Massachusetts.*

DEAR MR. SECRETARY: — I, Francis W. Sargent, pursuant to the provisions of Article XLVIII of the Amendments to the Constitution,

APPENDIX CPackaged Commodities Regulated and Unit of Measure

The following commodities shall be labeled in accordance with these Regulations. Each commodity must be unit priced only in the unit of measure given below.

<u>Commodity</u>	<u>Unit of Measure</u>
Aluminum and plastic wraps and waxed paper	100 square feet
Baby foods	Quarts if sold by volume Pounds if sold by weight
Baking mixes and supplies	Pounds
Bottled and canned beverages	Gallons
Bread and pastry products	Pounds
Candy in sizes greater than three ounces	Pounds
Canned poultry, fish and meat and poultry, fish and meat products	Pounds
Cereals, dry, ready-to-eat	Pounds
Cocoa	Pounds
Coffee - instant and ground	Pounds
Convenience dinners, "one-pan" meals	Pounds
Cookies and crackers	Pounds
Dairy products Gallons for: Milk, eggnog Quarts for cream and sour cream Yogurt	Pounds for: Cheese and cheese products
Deodorants	Pounds
Detergent and soaps	Pounds if sold by weight Gallons if sold by volume

Flour	Pounds
Fresh vegetables and produce	May be sold by volume, weight or per each PROVIDED that the total measured quantity, item price and unit price are clearly marked for every item. Individual items sold per each shall be exempt from these regulations.
Frozen foods except dairy products	Pounds
Fruits, vegetables, and juices canned, jarred, box, frozen and unfrozen	Pounds for all fruits and vegetables, except juices - quarts for all juices
Grains and beans	Pounds
Hair conditioners to include hair rinses and hair sprays	Pounds if sold by weight Quarts if sold by volume
Household cleansers, waxes, deodorizers, starches and bleaches	Pounds if sold by weight Gallons if sold by volume
Instant breakfast foods	Pounds
Jellies, jams, preserves, and sandwich spreads	Pounds
Ketchups, mustards, sauces, and condiments (including pickles and olives)	Pounds if sold by weight Quarts if sold by volume
Oleomargarine and butter	Pounds
Pet food, canned and dried, and kitty litter	Pounds
Plastic bags	100 count
Retail sales of food made from bulk if the quantity is weighed, measured or counted by the retailer, such as:	
Cold cuts	Pounds
Fish products and meat	Pounds

Salad dressings including mayonnaise	Quarts if sold by volume Pounds if sold by weight
Salads	Pounds
Salt	Pounds
Sanitary paper products, including but not limited to napkins, facial tissues, bathroom tissues Paper towels	100 count 100 square feet
Seasonings and spices in sizes greater than 3 ounces	Pounds
Shampoos	Quarts
Shaving cream	Pounds
Shortenings	Pounds if solid Gallons if liquid
Snack foods, including but not limited to potato chips, pretzels, and nuts	Pounds
Soups	Pounds if sold by weight
Sugar	Pounds
Syrups	Quarts
Tea	Pounds if sold loose Per 100 units if sold as tea bags
Toothpaste	Pounds





CONSUMER RESEARCH OF THE OFFICE OF ENERGY CONSERVATION  
AND ENVIRONMENT, FEDERAL ENERGY ADMINISTRATION

Jeffrey S. Milstein  
Federal Energy Administration

The purpose of the consumer research program of the Office of Energy Conservation and Environment, FEA, is to determine how to encourage a general conservation ethic as well as accelerate the utilization of energy-conserving goods and practices. Consumer researchers can contribute to this program by bidding on formal requests for proposals, by submitting research ideas for incorporation into future requests for proposals, and as consultants to help develop the consumer research program.

One of the main objectives of the energy conservation policies and programs of the Federal Energy Administration is to increase the energy-conserving behavior of the American consumer. FEA is trying to encourage a general conservation ethic as well as accelerate the utilization of specific energy-conserving goods (such as insulation, more efficient cars, and other consumer goods). The energy conservation ethic manifests itself in desirable energy-conserving behaviors, such as turning off unneeded lights or taking a bus instead of driving one's own car, and making more discriminating market demands for energy-saving consumer products. The purpose of the consumer research program of the Office of Energy Conservation and Environment is to determine for FEA how this popular encouragement can be done most effectively, to evaluate how effectively it is being done in terms of the end results of consumers' energy-conserving behavior, and to determine the factors in both the public and private sectors that facilitate and inhibit this process.

The methods and concepts of consumer research are used to determine in the most systematic and scientific way possible how energy scarcities, higher prices, and policies at all levels of government and business affect the American consumers' attitudes and actual energy-saving behavior. The results of these consumer research studies are used to develop and evaluate the effectiveness of energy conservation policies and programs at the Federal, state, and local levels.

I have conceptualized five general inter-related factors that impact on consumers' energy-conserving behavior, either directly or indirectly. These general factors are:

1. Energy supplies available to the consumer;
2. Government energy policies and programs;
3. Consumers' awareness, attitudes, and preferences about energy use and motivations to conserve energy;
4. Social and economic structure, including business policies and practices that provide energy-using goods and services to consumers; and
5. Consumers' actual energy-conserving and using behavior.

Consumer research is undertaken to obtain for FEA data on each of these factors, and analyses performed to show how these factors impact on each other. With the information resulting from this on-going and planned research, FEA will be better able to stimulate an optimum of energy-conserving behavior by the American consumer.

In general, the consumer research is divided into five major sets. The first is an analysis of how reductions in energy supplies and government energy policy and programs affect consumers' awareness, attitudes, and preferences regarding energy use and motivations to conserve energy and consumers' energy-conserving behavior.

To obtain data for this purpose, the Office of Energy Conservation and Environment of FEA has contracted out for a weekly national survey. The sample is a stratified random sample of 300 persons per week (a different sample is chosen each week). The questions we ask seek to determine attitudes and awareness towards energy policies, overall knowledge about energy and its use by consumers, factors that motivate people to use and save energy, what consumers' reported energy using behavior is, what effects energy shortages and/or higher prices have on their behavior, and their attitudes towards the energy shortage, higher energy prices, and the uses of energy.

Typical subjects we query about include the perceived seriousness of the energy shortage, perceived responsibility for energy shortages, how much consumers have been affected by energy shortages, the perceived role of the federal government, perceived reliability of information sources about energy, perception of actual or possible government policies, personal efforts to conserve energy, personal steps taken to use less energy (e.g. insulating home, driving slower, buying smaller cars, car pooling, using mass transit, turning down thermostats, using less lighting and air conditioning, changing shopping patterns, and traveling less). We also try to assess the knowledge and understanding of the people regarding balance of international trade, the impending coal strike, etc.

We use the results of these surveys to evaluate the effectiveness of voluntary conservation programs, and to evaluate policy options to deal with further conservation requirements.

The second research set is designed to find out what motivates consumers to use energy and to conserve energy, and then to determine what program of advertising and media communications (both persuasive and informational) would be most effective in getting consumers to conserve energy. The communications to consumers regarding energy conservation include a public service advertising campaign sponsored by the Advertising Council, with advertising in all the media. In addition, there are proposed or actual voluntary and/or mandatory labeling requirements on cars and some kinds of appliances. Further, there are now legal requirements that fuel oil distributors send all their customers information on how to conserve energy in heating their homes. There are also actual or proposed television programs on energy use and energy conservation.

We basically must find out how to communicate most effectively to get American consumers to conserve energy.

The third research set is designed to find out how business policies and practices affect consumers' energy-conservation attitudes and behavior through advertising and product lines.

The fourth research set is designed to develop a set of government policies and programs that will hasten the provision of energy-efficient products and services for the consumer and give those products (including energy-efficient housing and transportation) a competitive advantage in the market.

The fifth research set is designed to identify the economic and social consequences of reducing the energy demand growth rate and evaluating whether these consequences help or hinder energy conservation by consumers. Such consequences include the quality of life American consumers enjoy (absolutely and in comparison to wealthy foreign countries that consume less energy per capita than in the U. S.), and conflicts consumers might have with others in the society.

Now how can consumer researchers contribute to this consumer research program of FEA's Office of Energy Conservation and Environment? First, I must explain that all of our consumer research is done by people outside the Government who are contracted by FEA to do this research.

FEA is not a grant-giving organization: all of our contract research is done by competitive bidding in response to formal requests for proposals which FEA advertises in Commerce Business Daily. Those requests for proposals are related to our programmatic research which I have just outlined.

Unsolicited proposals for research cannot be funded on a sole-source contact basis. If one has an idea for research to be done, and is willing to compete to do the research, that idea may be incorporated into one of our requests for proposals and become a part of our programmatic research if communicated to us.

FEA occasionally also hires consultants on a daily basis to help develop our consumer research program. Such consultants, however, would be precluded from competing for a research contract on any research he or she had a hand in helping develop.

Finally, FEA welcomes the opportunity to communicate with consumer researchers formally and informally, to read their published and unpublished reports, in an overall attempt to help us to better do our job of getting the American consumer to conserve energy.



AUTHOR INDEX

Aaby, N.E.....	643	Hempel, D.J.....71,	231
Albaum, G.....	507	Hilliard, J.E.....	157
Arndt, J.....	63	Holbrook, M.B.....	375
Baron, P.....655,	679	Homan, G.....	217
Batsell, R.R.....	165	Huber, J.....	515
Battalio, R.C.....	657	Huff, D.L.....	165
Becker, H.....	141	Hulbert, J.M.....	375
Belk, R.W.....	427	Jain, S.C.....	39
Bettman, J.R.....	357	Johnson, K.....	203
Bhagat, R.S.....	405	Juster, F.T.....	725
Blackwell, R.D.....817,	837	Kagel, J.H.....	657
Bonfield, E.H.....	265	Kahn, F.....595,	751
Brooker, G.....	563	Kakkar, P.....	439
Bruno, A.V.....	757	Kassarjian, H.H.....	631
Burnkrant, R.E.....	465	Keuhl, P.G.....	51
Calder, B.J.....	241	Kotler, P.....	761
Capon, N.....	357	Leavitt, C.....545,	607
Carlson, C.J.....	631	Lehmann, D.R.....	741
Cecil, R.....	217	Light, L.....595,	751
Clokey, J.D.....	335	Lloyd, C.....	679
Combes, J.M.....	335	Lovelock, C.H.....	851
Craig, C.S.....	555	Lutz, R.J.....357,	439
Crane, E.....	63	Mason, J.B.....	173
Culley, J.D.....	879	McEwen, W.J.....	231
Cummings, W.H.....	21	McGann, A.....	643
Dacey, R.....	671	Miller, B.C.....	93
Deutscher, T.....	285	Miller, K.E.....	455
Dickson, J.....	507	Miller, S.J.....	187
Dubois, B.....	321	Milstein, J.S.....	925
Dunsing, M.M.....	103	Mitchell, A.A.....	249
Dyer, R.F.....	51	Mizerski, R.W.....	471
Ellis, R.B.....	481	Nimmo, D.....	771
Engledow, J.L.....	141	Nwokoye, N.G.....	693
Eskin, G.....655,	679	Oliver, R.L.....	295
Evans, R.H.....	525	Olson, D.W.....	495
Felson, M.....	33	Olson, J.C.....	249
Ferber, R.....	113	Omura, G.S.....	803
Ford, G.T.....	51	O'Neill, H.....	863
Gillette, W.....	525	Ott, L.E.....	589
Ginter, J.L.....	555	Parkinson, T.L.....	133
Haefner, J.E.....	905	Pessemier, E.A.....	705
Hafstrom, J.L.....	103	Peter, J.P.....	119
Hansen, F.....	307	Raju, P.S.....	405
Hawes, D.K.....	817		

(continued on next page)

AUTHOR INDEX  
(cont'd.)

Ray, M.L.....	577
Reynolds, F.D.....	157
Richard, L.M.....	535
Robinson, J.P.....	847
Rohloff, A.C.....	755
Rosin, P.E.....	631
Ross, I.....	1
Ryan, M.J.....	265
Ryans, A.B.....	285
Sawyer, A.G.....	713
Sexton, D.E.....	197
Shama, A.....	793
Sheth, J.N.....	389, 405
Stiff, R.....	203
Talarzyk, W.W.....	803, 817
Tarpey, L.X., Sr.....	119
Thorelli, H.B.....	141
Tourk, K.A.....	203
Tuncaip, S.....	389
Vaughn, R.L.....	157
Venkatesan, M.....	21
Voss, J.L.....	837
Walton, J.....	545
Waung, S.....	345
Wax, S.B.....	915
Wells, W.....	217
Wiley, J.B.....	535
Williams, D.A.....	615
Williams, F.....	891
Woodside, A.G.....	81, 335
Yalch, R.F.....	783
Zikmund, W.G.....	187