A NORMATIVE METHODOLOGY
FOR MODELING CONSUMER RESPONSE
TO INNOVATION

by

John R. Hauser[*] and Glen L. Urban[**]

Working Paper 854-76                    March 1976

* Assistant Professor of Marketing and Transportation,
  Graduate School of Management, Northwestern University

** Associate Professor of Management Science,
   Sloan School of Management, Massachusetts Institute of
   Technology

## Acknowledgements

# Abstract

Consumer response determines the success or failure of new products and
services. This paper proposes a methodology which integrates knowledge in the
fields of psychometrics, utility theory, and stochastic choice theory to improve
the design of new products and services. The methodology consists of a consumer
response and a managerial design process. The design process is one of idea
generation, evaluation, and refinement while the consumer response is based
on consumer measurement, models of the individual choice process, and aggregation
of predictions of indivdual choices. The individual response model processes
the consumer measures by first reducing them to an underlying set of perceptual
dimensions. Then the measures of perception are combined to produce a scalar
goodness measure for each choice alternative through a process called "compaction".
Next, homogeneous segments are defined based on similar preferences. The
goodness measures for each consumer or segment are linked to probability of
choice for the new products and services and for competing products and
services. In each step theoretical, empirical, and statistical issues are
identified. Various existing techniques and new techniques are introduced and
described for each phase. Selected techniques are demonstrated based on the
survey data collected at MIT to support the design of a health maintenance
organization (HMO) and in the consumer market to evaluate a new deodorant.

## 1. Introduction

Almost all organizations face the common problem of how to develop and introduce successful new products or services. This problem receives high strategic importance since such innovation is linked to increased effectiveness and productivity. In the private firm successful new products result in sales and profit growth. For example, approximately 50% of the growth in sales over a five year period in many industries were accounted for by new products (Booz, Allen, and Hamilton [5]). In services like transportation, additional ridership, efficiency, and revenue may be obtained by new services such as computer controlled mini-buses. Innovation in the design of service packages in the fields of insurance and finance can improve the competitive positions of companies and insure a stable base for corporate growth. In the field of health, the health maintenance organization (HMO) provides an example of an innovative new service. A successful HMO could have lower costs and higher quality of care along with high enrollment and re-enrollment. Although the measures of effectiveness vary across public and private industries, new products and services are critical to vital functioning and achievement of goals.

While new products and services are crucial to organizational growth and effectiveness, they also represent a high risk to the organization. Many new products fail. Approximately 30% of the new products introduced by firms in the market fail and 80% of the resources for new product development are allocated to products that are not a success in the market (Booz, Allen, and Hamilton [5]). In public organizations many failures have been recorded in public programs such as low cost housing, mass transit services, and preventive health services. Many of these failures reflect

a lack of acceptance by consumers. The products did not sell enough
or the public services were not utilized by the clientele.

The critical role of the consumer in the management of innovation is
being more clearly recognized. Private firms who sell directly to the mass
market have long recognized that understanding consumer needs is the key to
successful innovation. New evidence indicates that even in high technology
areas 60 to 80% of successful technological innovations are generated by
consideration of consumer needs (Utterback [66]). Since it is clear that most
successful innovation in private firms is due to understanding consumer needs
and responses, it is probably reasonable to posit that this same effort directed
at consumers could increase the rate of success of innovations in public
organizations as well.

Private firms allocate substantial effort and resources to developing new
products through R&D and marketing departments. Figure One depicts a process
for development of innovative products and services. The first step is design.
In this step consumer studies are integrated with technology and creative
efforts to generate new ideas. These ideas are then evaluated and refined
based on consumer reactions, production issues, and financial considerations.
After an idea has been established as a viable and significant innovation, it
is tested in a pilot program or test market. If the test is successful, the
product can be introduced.

This paper will address the problems of the design stage of new product
development. Emphasis will be on integrating consumer response into the
design activities of idea generation, evaluation, and refinement. This
integration will be done through a behavioral process model of individual
response to innovation. After summarizing the most relevant existing work in
the fields of psychometrics, utility theory, and stochastic choice theory,

the macro model structure will be defined. Next the measurement, estimation, and micro structural issues will be discussed. Specific examples will be provided based on: (1) the problem of designing a new prepaid, comprehensive health service plan (HMO) and (2) on consumer evaluation of a new deodorant. The paper will close with a description of future research needs.
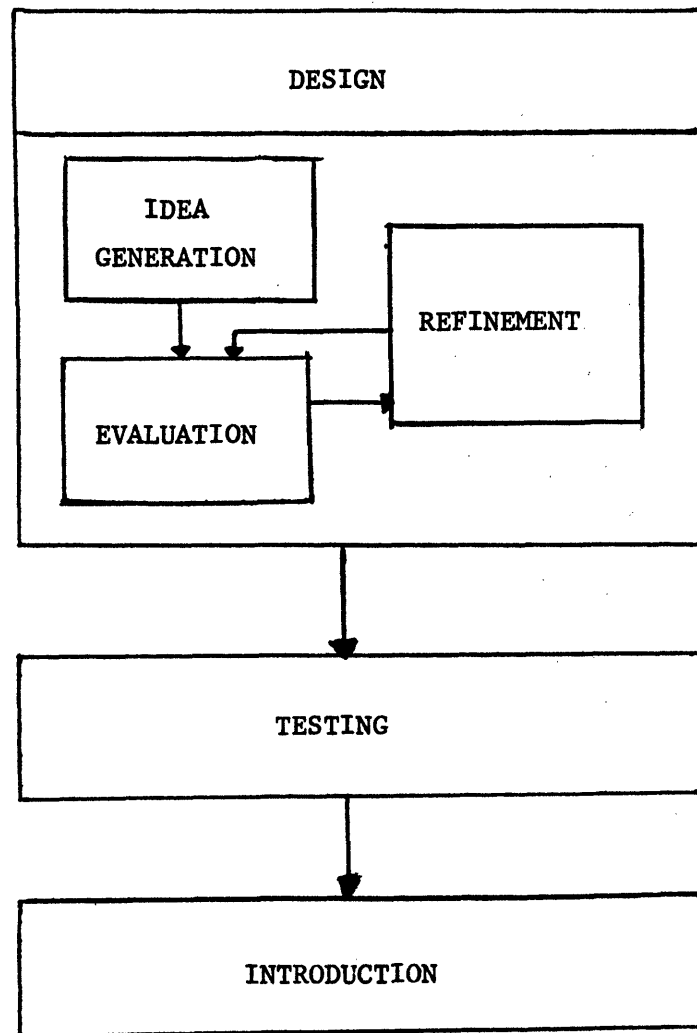
[Insert Figure One here.]

Figure 1: Process for Development of New
Products and Services

## 2. Some Existing Work

### Psychometrics

Psychometricians are concerned with the problem of how individuals perceive stimuli. Using measurements of perceived similarity among stimuli and measurements of attributes for new and existing stimuli, perceptual maps can be developed by multi-dimensional scaling procedures (Kruskal [38], Young and Torgerson [71]). These perceptual maps identify the important dimensions which consumers use to distinguish between stimuli and indicate the position of each stimuli relative to these dimensions. In marketing, the stimuli are products and the map defines market structure. Opportunities for new products are identified by examining the gaps in the market structure (Stefflre [59], Green and Carmone [16]).

Preference judgements can be integrated with the perceptual data to indicate high opportunity areas. PREFMAP is a popular method for accomplishing this task (Carroll and Chang [8], Carroll [6]). PREFMAP uses regression to derive an "ideal" point and relative importances of the dimensions from stated preferences of the consumers regarding the existing stimuli. Srinivasan and Shocker [58] have developed an alternative fitting procedure for estimating importances utilizing linear programming. Another approach is through conjoint analysis (Tversky [62]), which draws on an axiomatic and statistical base to produce relative importances by requiring consumers to rank order preferences for factorially generated combinations of product attributes (Green and Wind [20], Johnson [27]).

While PREFMAP and conjoint analysis use statistical procedures to impute the importances, other psychologists use "expectancy vlaue" models which utilize direct consumer judgements to estimate importances. Extensive work

has been done on such models based on psychological theories of attitude formation (Fishbein [11], Rosenburg [52]) and extensions on their work (Ryan and Bonfield [53]). Most of the models are conceptually similar in that they define an attitude towards an object as a linear additive function of an individual's reactions to an object on an attribute scale multiplied by a measure of the effect of that attribute in the overall attitude formation. A common formulation is the linear combination of the "importance" of each attribute multiplied by the individual's belief as to the extent to which the attribute is offered by a specific alternative (Wilkie and Pessemier [69]).

## Utility Theory

While the psychometricians apply a methodology based on multi-dimensional scaling and statistical preference analysis, utility theorists approach a similar problem from a substantially different point of view. Prescriptive utility theory is oriented towards helping managers make policy decisions under uncertainty and derives its strength from a rigorous set of axioms (von Neumann and Morgenstern [67]) and theorems which specify unique functional forms, e.g., additive, multiplicative, and quasi-additive (Raiffa and Schlaifer [50], and Keeney [30, 32, 33], Richard [51], Farquhar [10], and Fishburn [12]). The coefficients of these functions reflect the relative importances of the relevant performance measures, their interdependencies, and the risk averseness of the decision maker. The theorems also indicate techniques to directly assess and test the preference parameters by asking individuals to state when they are "indifferent" between two alternative stimuli. The dependent value of the utility function is a single cardinal measure of goodness of an alternative. Since the theory is used to guide the decision rather than describe

it, the decision maker chooses the alternative with the highest expected utility value.

Most empirical applications have been based on directly assessing the utility function of one or a small number of decision makers based on a set of quantifiable attributes of alternatives (Keeney [31]). This is in contrast to the psychometrician's approach which is based on interviewing many consumers based on perceived attributes that must be individually scaled.

## Stochastic Choice Theory

Recognizing that there will always be uncertainty in any prediction of choice behavior, economists, transportation demand theorists, and mathematical psychologists concentrate on axioms to determine selection probabilities from observable "scale" values (Luce [40], McFadden [41]). Economists and demand theorists parameterize scale functions and statistically estimate the parameters from observations on actual choice among existing alternatives. Popular models for this are the multinominal logit and other "random utility" models (McFadden [42]).

Mathematical sociologists model the stochastic choice process directly through diffusion, learning, Bernoulli and semi-Markov models (Massy, Montgomery, and Morrison [45]). These models describe the dynamics of choice probabilities over time but do not link attributes of products or consumer preferences to choice.

## Discussion

Although a good deal of work is being done, it is clear that the work is very diverse. Each discipline reflects different measurement approaches, analytic techniques, and foci. Psychometricians are concerned with perceived

attributes and recovery of importances from stated consumer perceptions and preferences. Utility theorists are concerned with theoretical soundness of functional forms through axiomatic consistency and with direct assessment of relative importances, attribute interdependence, and the risk characteristics for the purpose of aiding decision making. Choice theorists axiomatically model linkages to probability of choice, but do not consider linkages between consumer perception and managerial prediction of attributes or axiomatic specification of their functional utility forms.

The approaches also differ in how they treat the issues of aggregation. Most psychometricians develop average representations of perception and preference, but explicitly check that they are homogeneous with respect to perception and preference (Carroll and Chang [7], Tucker and Messick [61]). Utility theorists and conjoint analysts work completely idiosyncratically. Demand choice theorists can directly model individual response, but their statistical techniques force judgemental specification of aggregate segments before parameter estimation.

While the approaches are incomplete and diverse, they are complementary with each being primarily directed at a different phase in the consumer choice process. We visualize a process of perception, preference, and choice that integrates the approaches to form a complete consumer response model. Some initial work has been done to integrate these disciplines, but only at an aggregate level (Urban [64], Pessemier [46]).

## 3. Macro Description of the Methodology

This section proposes a methodology that draws on the existing work in psychometrics, utility theory, and stochastic modeling. It attempts to be comprehensive by integrating existing approaches into a cohesive but modular process which offers a variety of techniques of varying complexity and data requirements. It utilizes structures that reflect the acceptance phenomena at a level consistent with what is known about behavior. Efforts are made to make assumptions explicit, to isolate weaknesses in existing techniques, indicate where improvements need to be made, and prevent models from being used in applications which violate their assumptions. Attention is focussed on models that can predict response to changes in design and that can be extended to design changes or to new alternatives which are outside of existing consumer experience. For example, the models make predictions of consumer response to a new HMO, even though none currently exist in the community.

Since the ultimate value of the methodology will be in better design of products and services, creativity is recognized as a critical element. The methodology elicits and focuses creativity by identifying characteristics relevant to the choice process and by explicitly measuring relative importance of these characteristics. Although some steps are technically complex, the underlying choice process structure is understandable to non-technical as well as technical members of a product team. We attempt to make the outputs of each step clear and understandable so the design team can visualize the choice process and can create and refine new products or services.

Design Process

The methodology is shown graphically illustrated in Figure Two as a managerial design process and a parallel consumer response process. The analytics, and the focus of this paper are in the consumer response process. First, measures of perception and preference with respect to the relevant choice alternatives are observed for a sample of consumers. These measures are used to estimate the parameters of a model of the individual choice process. Finally, an estimate of group response is obtained by aggregating individual acceptance measures (probabilities). The measures of group response are then input to the evaluation model in the managerial design process which includes consideration of investment, operating costs, risk, and externalities as well as consumer acceptance. [Insert Figure Two here.]

It is rare that a new product or service will be implemented based upon a single cycle through the methodology. Instead a screening process will result which identifies the most promising alternatives for further consideration. These alternatives are refined based on detailed diagnostic information generated by the individual choice models (see arrow marked B in Figure 2). The refined design can be analytically tested in the individual choice models and the simulated results can be used iteratively to lead towards a "best" design (see Arrows A and B in Figure 2). This "best" design identified by the iterative process is then tested by taking new consumer measures and cycling through the entire methodology (arrows C and D). In early design phases attention is upon design specification and improvement (arrows A and B) while in later phases of design, attention is focussed on evaluation and refinement (arrows C and D).

In early phases of new product or service design measures can only be obtained on concept descriptions. As the product or service design evolves
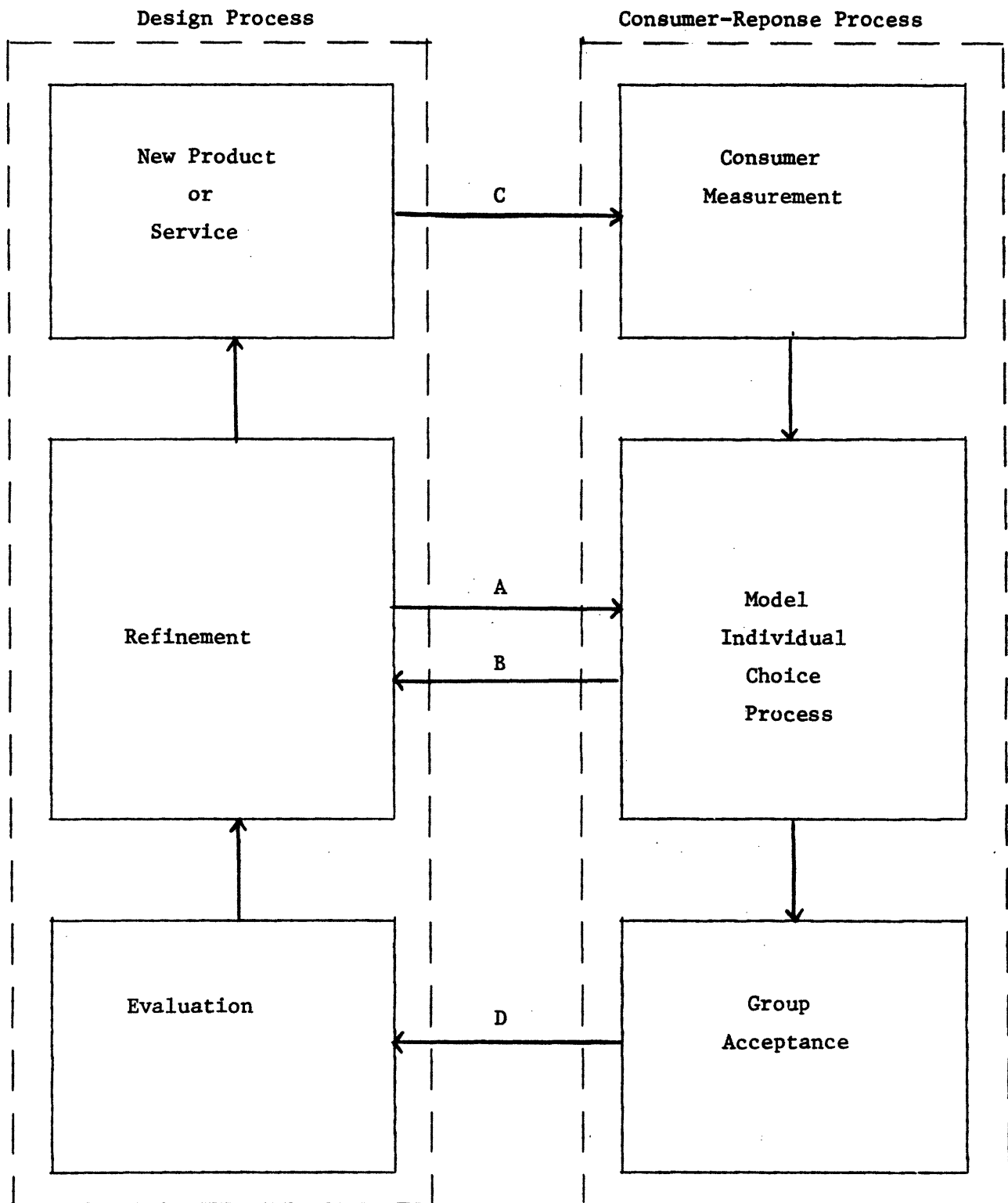
Figure 2:  Macro Description of the Methodology

more comprehensive descriptions become possible until in the final phases the product concept is a real choice alternative executed to the stage of advertising copy, package design, price, promotion, and distribution strategy. In this paper we present an early design problem for a new health service. Since additional models and measurement become necessary in later design phases, we illustrate these with an example from the evaluation of a new frequently purchased consumer product.

We begin our discussion with a brief overview of the consumer response model. The concepts, the models, the measures, and the notation will later be discussed in detail and illustrated with empirical examples.

## Consumer Response Components (Overview)

The basic input for the consumer response model is generated by surveys of the potential users of the new product or service alternatives. This set of information is denoted by $\Omega$. The individual choice process consists of the analytic phases of (1) perception, (2) "compaction", (3) segmentation, and (4) probability of choice. See Figure 3. The consumer response process is modular because this structure allows more effective integration of the disciplines of psychometrics, utility theory, and stochastic choice theory which are each closely associated with one module. In the perception phase the attitude evaluations of choice alternatives in measurement set ($\Omega$) are reduced to a smaller set of underlying perceptual performance dimensions. Emphasis is on the designation of the number of underlying dimensions and their names. This reduced set of perceptions is represented by X and is made up of each consumer's (i) perception of each performance dimension (m) for each choice alternative (j).

[Insert Figure three here.]

      Consumer Response Process

New Product or

Service

Refinement

Evaluation

**Consumer Measurement**

1. Choice Alternatives

2. Consumer Survey $(\Omega)$

**Individual Choice**

Perception

$\Omega \rightarrow X$

Compaction

$c_{ij} = c_s(\underline{x}_{ij}, \underline{\lambda}_i)$

Segmentation

s

Probability of Choice

$P_{ij} = p_s(a_i | c_{i1}, \ c_{i2} \cdots c_{ij} \cdots c_{iJ})$

Aggregation

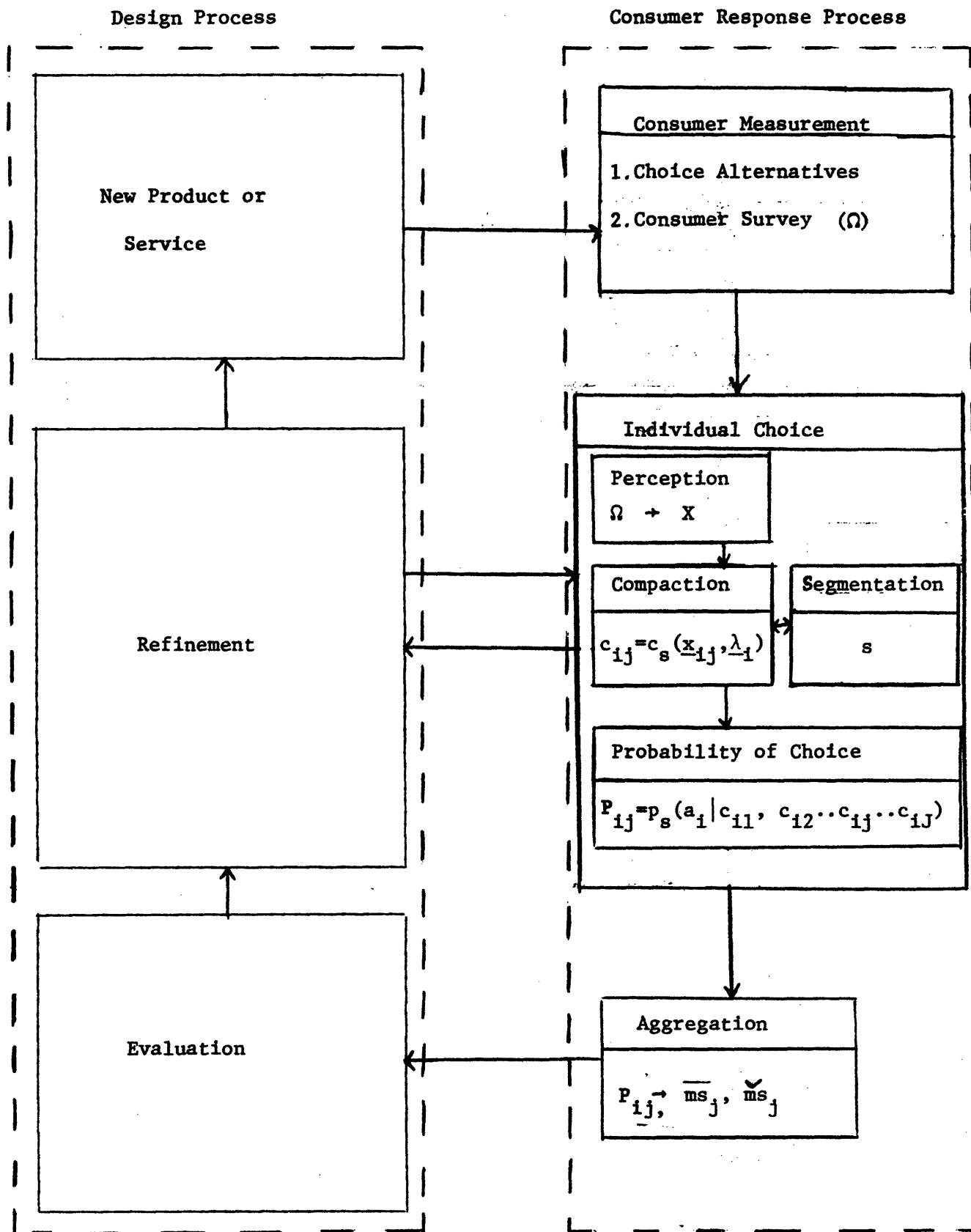$P_{ij} \rightarrow \overline{ms}_j, \ \breve{ms}_j$

Figure 3: Consumer Response Components

Next these multiple measures of perception (X) are combined to yield one measure of the goodness for each alternative. We choose to call this operation "compaction" because the several perceptual measures of an alternative are "compacted" into one measure of evaluation. We defined a new descriptive word since the procedures of utility theory, PREFMAP, expectancy value models, and conjoint analysis are all directed at this one task. In the compaction phase, the vector of individual performance scale values for each alternative for each dimension $(\underline{x}_{ij} = (x_{ij1}, x_{ij2}, \ldots x_{ijm} \ldots x_{ijM}))$ and a vector of individual parameters $(\underline{\lambda}_i)$ are variables in a real-valued function $(c(\underline{x}_{ij}, \underline{\lambda}_i))$ which compacts them into a scalar measure of goodness $(c_{ij})$. A separate goodness value is determined for each individual (i) in the sample and for each of his choice alternatives, $(a_j)$.

Based on the preference parameters $(\underline{\lambda}_i)$ of the compaction functions homogeneous groups of consumers are abstracted for designation as segments (s) of the target population. Within each segment (s) distributions of the performance dimensions $(X_s)$, the preference parameters $(\Lambda_s)$ and the functional form of the compaction function $(c_s(\underline{x}_{ij}, \underline{\lambda}_i))$ are determined. This segmentation is specific in its criteria of homoegeneity of the preference parameters $(\lambda_i)$ within the segment. For example, consumers who have similar importances for each perceptual dimension would be grouped together. Specific procedures and explicit tests for segmentation are discussed later in this paper and positioned versus existing segmentation methods.

Empirically it is observed that consumers do not always choose the alternative with the highest scalar measure of goodness. In this methodology, the scalar values are considered as independent variables in a probability of choice model which links an individual's vector of goodness measures

$(c_{11}, c_{12}, \ldots c_{ij}, \ldots c_{iJ})$ to his choice probabilities $(p_{ij})$. Each individual choice probability $(p_{ij})$ for each alternative $(a_j)$ is derived by a function $(p_s(a_j|c_{11}, c_{12}, \ldots c_{iJ}))$. The subscript $s$ indicates that the functions, but not the probabilities, are the same for all individuals in segment $s$.

The final step aggregates the individual choice probabilities to obtain group response measured by the mean $(\overline{ms}_j)$ and variance $(\overset{\vee}{ms}_j)$ of share of choice, or in some cases the mean $(\bar{N}_j)$ and variance $(\overset{\vee}{N}_j)$ of the total number of people choosing each alternative. If it is managerially useful, aggregation can be done separately within segments. By explicitly modeling the various stages of consumer response useful diagnostics can be identified for the refinement model. These diagnostics help the managers understand the market structure, segmentation, relative importances and interdependence of product attributes, and the risk characteristics of the choice process. The model itself processes individual information, but aggregate results from perception, segmentation, and compaction are important in eliciting creativity for the design of innovative alternatives. For example, in the perception phase, average perceptual maps of product positioning and in compaction, average values of the importance weights guide the manager in effective design of new products or services. These are particularly useful in the early design phase where the product must be positioned in the space of perceived product attributes. In the late design phase, such positioning is less of a concern and attention focuses on probability of choice and accurate forecasting and evaluation.

## 4. Micro Aspects of the Methodology

This section investigates each module of the consumer response process and gives examples of the measurement and estimation procedures. The major source of the examples is a study of health care innovation, but because the health care study (HMO study) concentrates on the early design process, some of the modules are futher illastrated with measures from the study of a new deodorant. Figure 4 summarizes the empirical examples. It is given now to enable the reader to view each example in its relationship to the total design process. Various techniques will be discussed throughout the paper, but all the in depth empirical examples are shown in Figure 4. In selecting empirical examples, we placed emphasis on new methods and techniques. Readers are referred to existing sources for well established conventional techniques.

[Insert Figure 4 here.]

### Consumer Measurement

A good model is dependent upon high quality input. A model is accurate only if the measurements it requires are valid and reliable. This section discusses the measurement issues of the methodology. The HMO measurement issues will be presented in some detail and then the deodorant data will be briefly discussed.

Early Design - HMO: Two samples were drawn from the target population for a new health maintenance organization at MIT. The first sample included faculty, students, and staff and was used for the perception, statistical compaction, and segmentation phases of the study. The second sample was drawn from the student subset of the target population and was used to test the

<u>Consumer Measurement</u>

<u>Choice Alternatives</u>

Early Design - HMO
        - existing health service
        - 3 concept descriptions

Late Design - Deodorant
        - existing evoked brands
        - complete new product execution

<u>Consumer Survey</u>

Early Design - HMO
        - campus wide sample
        - student sample - utility

Late Design - Deodorant
        - laboratory and exposure and choice
        - home use

<u>Individual Choice Model</u>

<u>Perception - HMO</u>

Factor analysis

<u>Compaction - HMO</u>

Preference regression - campus
    wide sample
Utility Theory - students
Saved Data Testing

<u>Segmentation</u>

Search for segments
Test of segmentation

<u>Probability of Choice</u>
Early Design - HMO
    - intent  & preference conversion
Late Design - Deodorant
    - Multinomial Logit
    - Testing of choice model

<u>Aggregation</u>
HMO - summation of choices
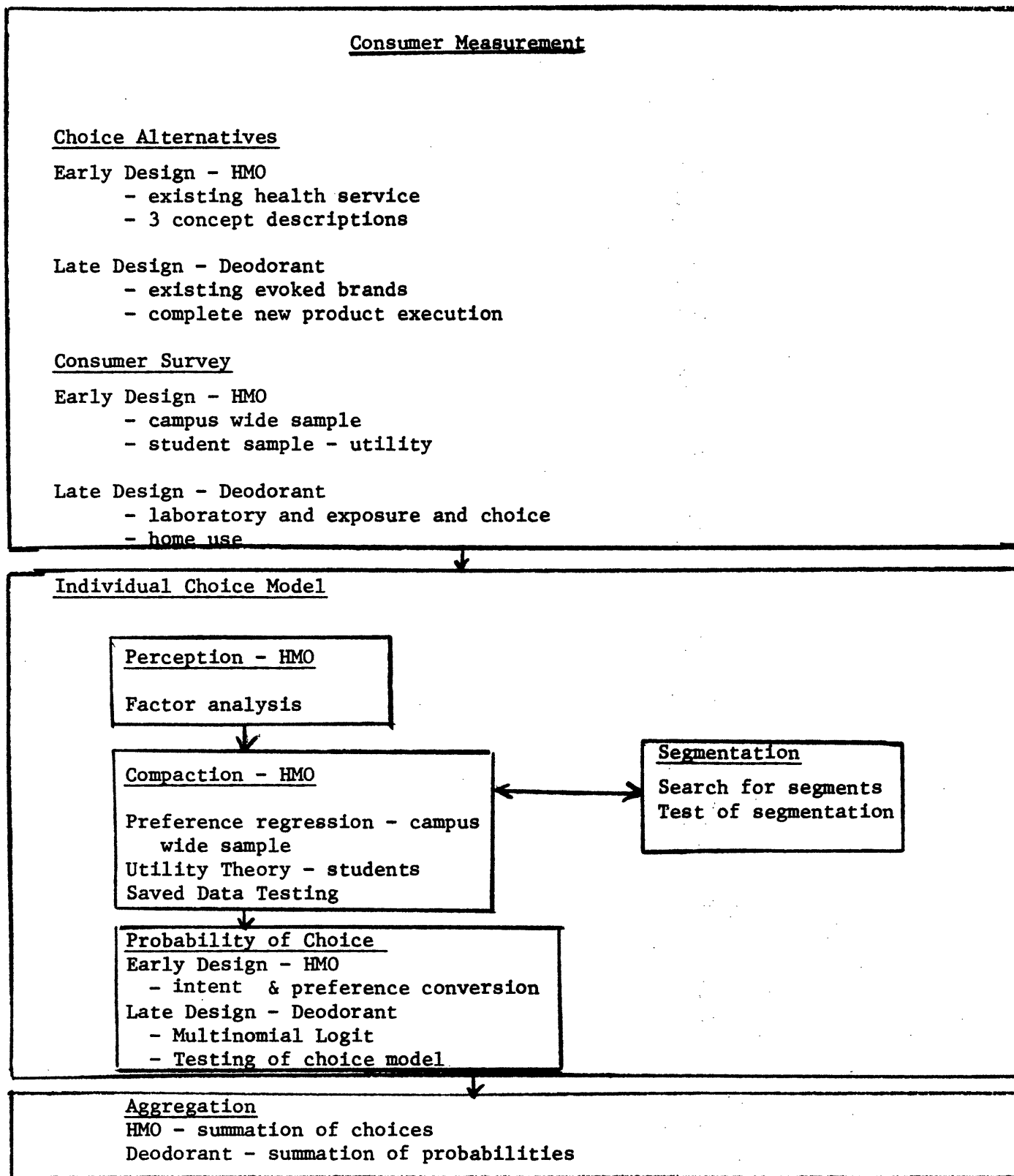Deodorant - summation of probabilities

Figure 4:  Guide to Empirical Presentation
          of Consumer Response Process

feasibility of applying von Neumann-Morgenstern utility theory for directly measuring compaction functions.

Consumer perceptions, preferences, and choice behavior must be observed with respect to the product or service alternatives which are "evoked". For example, there may be a large number of products available, but each consumer only "evokes" a few of these. In a study of seven consumer products, consumers had an average of only three "evoked" brands. "Evoked" was defined as brands last used, ever used, on hand, or would not consider using (Urban [64]). In services, especially new services, the number of alternatives is often so small that one must force evoking by the use of concept statements in order to have sufficient perceptual inputs. For example, in the study of HMO design at MIT the only real option available was existing private care. Thus the evoked set was expanded to four options by specifying three new options in concept form: an MIT HMO, the Harvard Community Health Plan, and a hypothetical Massachusetts Health Foundation. See Table 1 for an example of the MIT HMO option. One thousand surveys were mailed to a random sample of the MIT community and 447 faculty, students, and staff completed and returned the questionnaire.

[Insert Table One here.]

First a set of important product attributes are identified and consumers evaluate each evoked choice alternative with respect to each attribute. These attributes and their descriptions are generated from in-depth interviews with individual or groups of consumers or by Kelly's triad procedure (Kelly [34]), in which consumers describe how the two most similar of three stimuli are alike and how the two most dissimilar are different. Consumers then rate the attributes on bipolar or agree/disagree scales. In the HMO study, 16 statements consumers had earlier defined as relevant to their health care were evaluated by 5 point agree/disagree scales (see Table 2).

[Insert Table Two here.]

# DESCRIPTION OF THE M.I.T. HEALTH PLAN

M.I.T. announces a new health care plan for YOU AND YOUR FAMILY.  By joining the M.I.T. HEALTH PLAN you can get comprehensive health care at a low, fixed monthly charge.  Virtually all your medical needs will be met.  You will not have to face unexpected doctor or hospital bills and you will not have to worry about finding a good doctor for you or your family.

The cost of joining the M.I.T. HEALTH PLAN is only a little more than regular Blue Cross/Blue Shield health insurance, but you get more services and comprehensive care.  There are no charges for doctor visits, nursing and laboratory services, or hospital services.  Women in the plan pay nothing extra for prenatal, delivery, or maternity care.  The services are comprehensive and include mental health care and emergency services.

The costs are kept low by the utilization of preventive care to keep you well.  The plan succeeds by keeping you and your family well and out of the hospital.  In addition, the use of trained paramedicals and technology helps reduce costs while maintaining the quality of care.

You choose your own personal doctor (specialist in internal medicine for yourself and a pediatrician for your children) from our staff of physicians.  Your doctor supervises your total health care at the health center and in the hospital.  He will be sure you get the highest quality of care.  When you are a member of the M.I.T. HEALTH PLAN you can be sure of getting health care around the clock from the staff of physicians, nurses, social workers and allied health personnel.

The M.I.T. HEALTH PLAN delivers its services from the Homberg Memorial Building on the M.I.T campus.  Parking is available during patient visits.  Hospital services are provided by the Mount Auburn and Cambridge City Hospitals.  Maternity and gynecology care are provided through the resources of the Boston Hospital for Women.  For emergencies outside the Boston area, local hospitals can be used.

You can become a member of the plan by paying $1.50 per month more than your Blue Cross/Blue Shield coverage if you are single and $4.00 more per month if you are married.  If you are a single student and do <u>not</u> have hospital insurance, the cost is $8.25/month more than the student health fee you are currently paying; if you are a married student, the cost is $20.00/month more than the student health fee.  These fees cover all of your medical costs except: the first $50 and 20% of the balance of prescription charges and the excess of $10 per visit for psychotherapy (over $5 per visit for group therapy).  The plan does not include eye glasses, hearing aids, cosmetic surgery, custodial treatment, or dental care done outside a hospital.  If you join the plan, you must remain a member for one year.

The M.I.T. HEALTH PLAN is designed to make comprehensive, high quality health care available to you and your family at a low cost.

Table 1:  HMO Concept Description

| | strongly agree | agree | neither agree nor disagree | disagree | strongly disagree |
|---|---|---|---|---|---|
| 1. I would be able to get medical service and advice easily any time of the day and night. | a | b | c | d | e |
| 2. I would have to wait a long time to get service. | a | b | c | d | e |
| 3. I could trust that I am getting really good medical care. | a | b | c | d | e |
| 4. The health services would be inconveniently located and would be difficult to get to. | a | b | c | d | e |
| 5. I would be paying too much for my required medical services. | a | b | c | d | e |
| 6. I would get a friendly, warm and personal approach to my medical problems. | a | b | c | d | e |
| 7. The plan would help me prevent medical problems before they occurred. | a | b | c | d | e |
| 8. I could easily find a good doctor. | a | b | c | d | e |
| 9. The service would use modern, up-to-date treatment methods. | a | b | c | d | e |
| 10. No one has access to my medical record except medical personnel. | a | b | c | d | e |
| 11. There would not be a high continuing interest in my health care. | a | b | c | d | e |
| 12. The services would use the best possible hospitals. | a | b | c | d | e |
| 13. Too much work would be done by nurses and assistants rather than doctors. | a | b | c | d | e |
| 14. It would be an organized and complete medical service for me and my family. | a | b | c | d | e |
| 15. There would be much redtape and bureaucratic hassle. | a | b | c | d | e |
| 16. Highly competent doctors and specialists would be available to serve me. | a | b | c | d | e |
| (values assigned = | -2 | -1 | 0 | +1 | +2 ) |

Table 2:  Measuring Perceptions

After perceptions have been determined, rank order or constant sum preference
measures are obtained for use in statistical compaction techniques. Where
possible constant sum paired comparisons are preferred, since they yield
interval scales (Torgenson [60]). The initial HMO survey was done by mail and
hence the easier to answer rank order preference measures were collected.
Consumers were then asked to choose among alternatives. In the HMO study,
consumers were given choices (0, 1) between existing care and the HMO. If they
chose MIT, a 5-point intent scale was administered. Then other alternatives
were added to the choice set and new intent measures taken.

While the procedures for measurement of perception and preference are
comparatively well developed, this is not true in utility theory. There are
no reports of measurement of consumer utility functions. This is in part
because past uses of utility assessment were oriented as prescriptive decision
applications with one or few decision makers (Keeney [31]). When consumers
are considered, two assessment issues must be addressed:

(1) How can perceptual phenomena be integrated in the measurement? and

(2) Is it possible to have consumers understand the required lottery

and trade-off questions and give meaningful answers?

When directly assessed utility models are to be supported (see compaction
section of this paper for a more detailed description of the procedure), data
is required to measure risk averseness, importances, and interactions relative
to various "performance measures". Prescriptive utility theory requires
these to be quantifiable, instrumental variables such as cost or waiting time
rather than perceptual measures like quality or personalness of health care.
The former are easier for the manager to relate to, but the latter better
reflect the consumer choice process. In this methodology it is proposed that
the psychological dimensions obtained by reducing the perceptions of choice

alternatives be used as performance measures.

In the HMO case, a special study was conducted to assess the utility functions of 80 students. First 16 attributes were factor analyzed to obtain four perceptual measures (see reduction section of this paper for a more detailed description of the procedures). The dimensions were named "quality", "personalness", "convenience", and "value" and the 80 students additionally rated the same four health alternatives directly on 7-point scales for these four reduced perceptual dimensions. Utility functions were then assessed relative to the performance measures defined by the perceptions of the 7-point scales. After utility assessment the overall 7-point scale values can be correlated to the factor scores and therefore to the original perceptual ratings. By defining performance measures through perceptual scaling methods, utility theory can be meaningfully linked to psychometrics.

The use of a reduced number of perceptual dimensions as performance measures also makes consumer utility measurement more feasible. Since the number of dimensions is small (usually four or less), measures of risk aversion and importance only need be collected on this smaller set of performance measures.

Risk averseness is measured by having the consumer consider a lottery on the performance measures. Although it was anticipated that this would be a difficult task for consumers, it was found in the pilot study of 80 students, that they related well to a carefully designed questionnaire if simple lottery questions were included to educate them to the task required and the meaning of a probability. In fact, the only difficulty was with students already well schooled in probability theory who tried to give expected value answers rather than their true feelings. The procedure is schematically represented in Table 3. The respondent sets the area of a probability wheel so he would be

Instruction to Consumer:

Imagine you can only choose between two health plans, plan 1 and plan 2. In both plans personalness, convenience, and value are good (rated 5). You are familiar with plan 1 and know that the quality of plan 1 is satisfactory plus (rated 4). You are not sure of the quality of plan 2. If you choose plan 2, then the wheel is spun and the quality you will experience for the entire year depends on the outcome of the wheel. If it comes up yellow, the quality is very good (rated 6) and if it comes up blue the quality is just adequate (rated 2). Graphically this is stated:

```
                                          +----------------------------------+
                                          |            Plan 2                 |
                                          | Personalness  |                   |
                                          |  5 (Good)     |                   |
                                          |               |    Quality        |
                                          | Convenience   |      6            |
                                          |  5 (Good)     |  (Very Good)      |
                                          |               |                   |
                                          | Value         |                   |
                                          |  5 (Good)     |                   |
                                          +----------------------------------+
                                                  (Yellow Card)


+--------------------------------+
|           Plan 1               |        +----------------------------------+
| Personalness  |                |        |            Plan 2                 |
|  5 (Good)     |                |        | Personalness  |                   |
|               |   Quality      |        |  5 (Good)     |                   |
| Convenience   |     4          |        |               |    Quality        |
|  5 (Good)     | (Satisfactory  |        | Convenience   |      2            |
|               |    plus)       |        |  5 (Good)     | (Just adequate)   |
| Value         |                |        |               |                   |
|  5 (Good)     |                |        | Value         |                   |
+--------------------------------+        |  5 (Good)     |                   |
         (Green Card)                     +----------------------------------+
                                                   (Blue Card)
```

Yellow
Blue

+------------------------------------+
|              RULES                 |
|                                    |
|  - wheel is spun after you         |
|    make your decision              |
|                                    |
|  - you must accept the conse-      |
|    quences and cannot switch       |
+------------------------------------+

Instruction to Consumer:

At what setting of the odds (size of the yellow area) would you be indifferent between plan 1 and plan 2? (Respondent is given wheel and adjusts it until size of yellow area is appropriate. He is challenged by being given the choice with his setting. If he prefers one plan or the other, the interviewer iterates the question until a true indifference setting is determined.)

Table 3:  Schematic of Risk Aversion Question

indifferent between the certain health plan outcomes and uncertain health plans as represented by lottery outcomes. Most respondents were comfortable with this task and all completed the interview. Utility assessment requires two lotteries for each performance measure or a total of eight lottery judgements in the HMO case. One lottery, where one performance varies while all others are held fixed, is enough to determine the risk characteristics of that performance measure if constant risk aversion and utility independence are assumed (Raiffa [49], Keeney [30]). The second lottery is needed to verify behavioral assumptions inherent in the form of the utility function.

[Insert Table 3 here.]

Relative importance weights are determined by asking consumers to trade-off one performance measure, say convenience, with another, say value, while holding all others fixed. See Table 4. Another trade-off question then varies the fixed values to verify a behavioral assumption known as preferential independence, which together with utility independence, specifies uniquely the

[Insert Table 4 here.]

form of the utility function (Keeney [30]). Thus two trade-off questions are required for all but one of the performance measures. This makes a total of six trade-off questions and eight unidimensional lottery questions in the HMO case. Finally one complex lottery involving simultaneous changes in all the performance measures is used to measure interactions.

The results from the HMO study indicate perceptual dimensions can function adequately as performance measures and it is feasible at least in the case of students to measure consumer risk aversion and importances. Future research will assess the practicality of measuring utility parameters in a general population sample.

Instruction to Consumers:


      Now consider the two plans below and choose the level of the quality factor in such a way that you are indifferent between the two plans. (Consumer is challenged and the question iterated until a true indifference is determined).


Plan A

Value          = 5 (good)

Personalness   = 5 (good)

Convenience    = 6 (very good)

Quality       = 2 (very poor)

Plan B

Value          = 5 (good)

Personalness   = 5 (good)

Convenience    = 2 (very poor)

Quality = __ __ __ __ __ __ __

| Very Poor | Just adequate | Satisfactory | Satisfactory Plus | Good | Very good | Excellent |
|---|---|---|---|---|---|---|


Table 4:  Schematic of Trade-off Question

In both the campus wide and student utility sample demographics and other consumer descriptors were collected in order to adequately project from a sample population. For example, in the study, patterns of health care utilization and satisfaction were measured in addition to demographics such as age, sex, family size, and health status.

Late Design - Deodorant: As the product or service design is finalized the accurate forecasting of demand becomes more important and measurements are changed to more closely simulate actual environments. In the deodorant study, consumers were intercepted in a shopping mall, exposed to TV advertisements for the new and old products, given the opportunity to buy the product from a retail shelf, and took the product home for use. The reader is referred to Silk and Urban [56] for a detailed discussion of the measurement design and execution.

Since a personal interview was used, constant sum preference (Torgenson [60]) measures before exposure to the new product and after home use of the brand were collected. These before and after preference measures became the input to the probability of choice model which will be discussed later in this paper.

Perception

In the perception phase of the consumer response model, the attitude data collected in the measurement phase are reduced to a smaller underlying set of psychological dimensions. For completeness it is necessary to measure consumers on a large number of possible perceptual dimensions. But it is often difficult for managers and analysts to gain insight from comparison of

perceptions relative to a large number of scales. For example, Table 5 shows

an average rating of consumers for the four HMO alternatives on the 16 scales

shown in Table 2. What the manager requires is a simpler representation that

can be easily visualized and internalized for further processing. Thus in

the reduction module of the consumer response model, the perceptual data are

reduced to a smaller underlying set of psychological dimensions. These

dimensions capture the essense of the perceptual process in a form that is

readily understandable and more appropriate for use in design.

[Insert Table 5 here.]

Several multidimensional approaches were available. If similarity

judgements as well as ratings are collected, non-metric techniques can be used

to place the stimuli in perceptual space (Green and Carmone [16]). In many

studies the evoked set is too small ($n \leq 8$) or too varied across individuals

to use non-metric techniques to achieve statistical significance (Khlar [35]).

To overcome this, the ratings can be directly reduced by using factor analysis

on a data matrix in which each row reflects an individual's rating of a stimuli.

In this manner even if each individual only evokes a small number of alterna-

tives (e.g., $n \leq 4$) the number of observations is large and equal to the number

of individuals times the average number of choices evoked.

In all reduction methods care must be taken to test the results. In

this methodology the sufficiency of the reduction is tested by correlations

to preference and choice at later stages in the methodology. If the number of

dimensions and their interpretation is inappropriate, the compaction phase

preference prediction will be poor, and provide a warning to the analyst. In

the HMO study, principal component factor analysis was used to reduce the

ratings on the sixteen scales for the four plans across 234 individuals to four

underlying dimensions. These four factors explained 55 percent of the total

.8  .6  .2  0  .2  .4  .6  .8  1.0

Existing  HCHP  MIT

1.  I would be able to get medical
    service and advice easily any
    time of the day and night.

2.  I would have to wait a long time
    to get service.

3.  I could trust that I am getting
    really good medical care.

4.  The health services would be
    inconveniently located and would
    be difficult to get to.

5.  I would be paying too much for
    my required medical services.

6.  I would get a friendly, warm
    and personal approach to my
    medical problems.

7.  The plan would help me prevent
    medical problems before they
    occurred.

8.  I could easily find a good
    doctor.

9.  The service would use modern,
    up-to-date treatment methods.

10. No one has access to my medical
    record except medical personnel.

11. There would not be a high con-
    tinuing interest in my health
    care.

12. The services would use the best
    possible hospitals.

13. Too much work would be done by
    nurses and assistants rather
    than doctors.

14. It would be an organized and
    complete medical service for me
    and my family.

15. There would be much redtape and
    bureaucratic hassle.

16. Highly competent doctors and
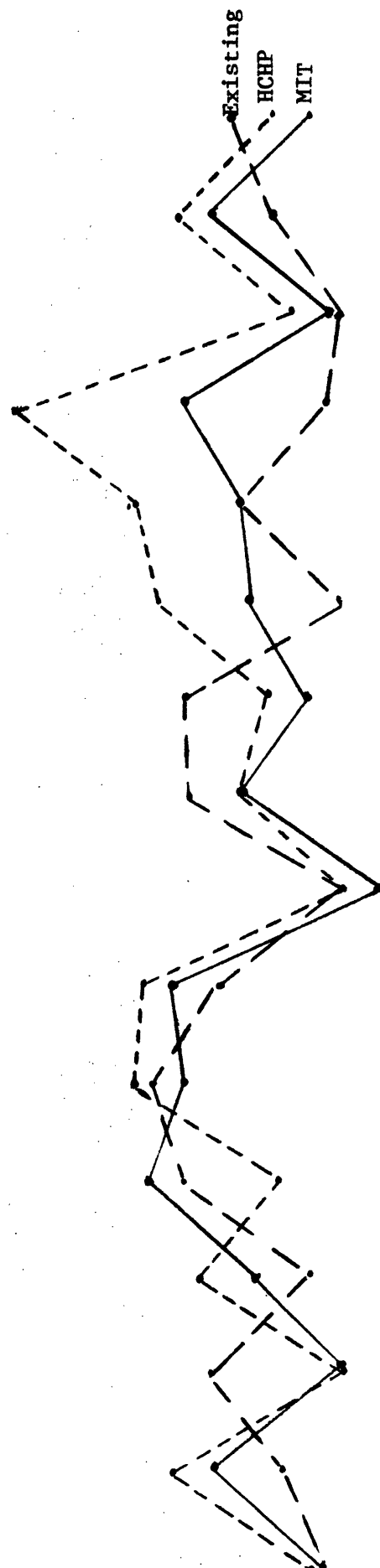    specialists would be available
    to serve me.

Table 5:  Average Ratings

variance. Table 6 presents the factor loadings (correlations) of the raw scales to the new underlying dimensions. By examining the high loadings on each dimension they were labeled judgmentally: (1) quality, (2) personalness, (3) value, and (4) convenience. Quality correlated to trust, preventive care, availability of good doctors, and hospitals. Personalness reflected a friendly atmosphere with privacy and no bureaucratic hassle. Value was not just price, but rather paying the right amount for the services. Convenience reflected location, waiting time, and hours of operation. The interpretation was similar based on a common factor analysis. Factor scores were obtained which described the location of each plan on each dimension for each individual.

[Insert Table 6 here.]

## Compaction

The first analytic module in the methodology tells us how consumers perceive the alternative products or services, but it does not tell us what tradeoffs consumers make in their decision to buy a product or select a service. To guide managerial decisions, the design team needs to know how consumers combine their perceptions on each performance dimension to form an evaluation of a product or service. For example, in an HMO should one increase the quality of the health service and charge a premium price or should the price be minimized subject to an adequate level of quality? Compaction generates this understanding by explicitly identifying the importance of each performance measure and their interaction in the consumers' evaluation of an alternative. Formally this means that the performance measures identified in reduction are now compacted to form for each individual a scalar measure of goodness for each alternative. Preference and ultimately probability of choice result from comparison of these goodness measures.

| ATTRIBUTE SCALE* | QUALITY | PERSONAL | VALUE | CONVENIENCE |
|---|---|---|---|---|
| 1. DAY & NIGHT CARE | 0.37244 | 0.07363 | -0.31379 | 0.63939 |
| 2. WAITING TIME | -0.22082 | 0.26204 | 0.15514 | -0.64870 |
| 3. TRUST-GOOD CARE | 0.72125 | -0.21828 | -0.09556 | 0.24708 |
| 4. LOCATION | 0.01144 | 0.24706 | -0.12544 | -0.72964 |
| 5. PRICE/VALUE | 0.03066 | 0.12810 | 0.72884 | -0.09961 |
| 6. FRIENDLY/PERSONAL | 0.40986 | -0.51317 | -0.12285 | 0.18768 |
| 7. PREVENTIVE CARE | 0.55403 | -0.14187 | -0.44353 | -0.03653 |
| 8. EASILY FIND GOOD MD | 0.64412 | -0.15036 | -0.21491 | 0.27113 |
| 9. MODERN TREATMENT | 0.72288 | -0.13441 | -0.15906 | 0.08018 |
| 10. ACCESS TO RECORDS | 0.43412 | -0.49053 | 0.18749 | -0.05992 |
| 11. CONTINUITY OF CARE | 0.20491 | 0.47900 | 0.47727 | 0.04725 |
| 12. ASSOCIATED HOSPITALS | 0.68006 | -0.08256 | 0.10854 | 0.00555 |
| 13. USE OF PARAMEDICALS | -0.05303 | 0.67083 | 0.12288 | 0.16722 |
| 14. ORGANIZED/COMPLETE | 0.47725 | 0.01627 | -0.52893 | 0.14816 |
| 15. HASSLE/REDTAPE | -0.13081 | 0.69824 | 0.11180 | -0.27903 |
| 16. COMPETENT MD'S | 0.73953 | -0.19335 | -0.13971 | 0.18691 |
| | | | | |
| EIGENVALUES | 5.34 | 1.4 | 1.1 | 1.02 |
| CUMULATIVE VARIANCE | .33 | .42 | .49 | .55 |

\* See Table 2 for field rating scale descriptions

Table 6: Factor Loadings

Specifically, in compaction a function, $c_s(\underline{x}_{ij}, \underline{\lambda}_i)$, is determined which

maps the vector of performance measures, $\underline{x}_{ij}$, into a scalar measure of goodness

(a real number). The performance measures are the result of the reduction

step in the methodology and the choice parameters result from the measured

preferences or tradeoffs and lotteries. For a given alternative, e.g., health

plan $a_j$, this scalar measure of goodness, $c_{ij}$, has the property that with

all other alternatives held fixed, any set of performance measures yielding

the same value, $c_{ij}$, must also yield the same probability of choice for

alternative $a_j$. In other words, compaction compresses the performance

measures for an alternative into a one-dimensional measure, and knowing the

value of this measure for each and every alternative is then sufficient to

predict choice.

There are a number of techniques to estimate the parameters of a compaction

function, but there are four basic categories: (1) direct consumer statement

of importances, (2) statistical estimation of choice parameters $(\underline{\lambda}_i)$,

(3) conjoint analysis, and (4) von Neumann-Morgenstern utility theory.

The most elementary technique is to ask consumers to scale the importance

of each measure and form a compaction function as a linear additive function

of an individual's reactions to an object on an attribute scale (e.g. rating

on scale) multiplied by a measure of the effect (e.g., importance) of that

attribute in the overall attitude formation. These models received considerable

attention in marketing from psychologists (see Wilkie and Pessemier [69] for

summary). Empirically, these models have been tested by correlation of the

predicted attitude value with preference or choice. Empirical results have

been mixed. Ryan and Bonfield [53] report correlations as high as .7 to .8 for

an extended Fishbein model while Seth and Talarzyk [54] report correlations in the

range of .1 to .4. Wilkie and Pessemier [69] in their review of 42 studies

identify 19 with favorable results, 14 questioning the model and 9 not applicable.

Another approach to compaction is to statistically estimate importances by a regression of observed preference against the perceptual attribute measures. PREFMAP (Carroll [6], Green and Carmone [16], Green and Rao [18]) is based on a regression of individual preference parameters for the consumers' evoked set of alternatives. In cases where the evoked set is small ($n \leq 8$) PREFMAP has very few degrees of freedom to estimate individual parameters. PREFMAP also estimates an "average ideal point", but this is based on the average preference across individuals and is subject to the same degrees of freedom limitations. Furthermore, it is a questionable procedure if evoked sets vary across respondents. Urban [64] has proposed an extension to the regression approach by grouping respondents and regressing across individuals <u>and</u> choice alternatives. This provides many degrees of freedom (number of individuals times average evoked set) and allows for individuals with differing evoked sets, but assumes the group is homogeneous with respect to the importance parameters ($\underline{\lambda}_i$). Grouping procedures that identify homogeneity will be more fully discussed in the segmentation section of this paper.

This extended preference regression method was applied to the campus-wide HMO data base of 234 individuals. A saved data same of 61 individuals was retained for predictive testing. The results are presented in Table 7. The regression was done across individuals and stimuli and there were 642 total observations. The regression was significant at the 1% level and all regression coefficients were significant at the 5% level. The $R^2$ fit statistic is not the most appropriate measure of fit since the dependent variable is rank ordered. Table 7 also reports a more appropriate measure — the fraction of times the predicted rank order preference was equal to the actual rank order preference. The observed fit of .44 can be compared to a random fit of .25. The Chi square statistic for the matrix of actual versus

predicted rank order was significant at the 5% level. In addition, the fraction of times first preference was indicated was also reported as .48 (random fraction is .25). These statistical results are quite encouraging and reflect adequate accuracy for compaction. Examination of the residuals indicated a unimodal distribution with 68% of the observations between $\pm 1\sigma$. The distribution was non-normal by the $\chi^2$ test ($X^2 = 50.2$ (df=8)) and the departure was largely due to fluctuation in the tails. Use of a non-linear form (logs of independent and dependent variables) and use of interaction terms did not improve the fits significantly.

[Insert Table 7 here.]

Since errors were not normally distributed and the dependent observations in the regression were individual rank order preference judgments, the assumption of an interval scale required in linear regression may not be appropriate. To test this, a monotonic regression was conducted with Johnson's monotonic regression program (Johnson [28]). Although the fit is improved, the normalized importances are very similar to the linear case. In the monotonic case, the importance of personalness is higher and value lower than in the linear case.

The estimated coefficients were used to predict the preference rank order for the 61 individuals in the saved data sample. The first preference was correctly predicted 36 percept of the time and the rank order 40 percent. The monotonic fit was again slightly better. The saved data fits are slightly lower than the estimated results, as expected, but the drop is not alarming and the empirical results are supported by saved data testing.

While the statistical approach to compaction is encouraging, there are disadvantages to this approach. To gain degrees of freedom consumers must be grouped prior to estimation. Furthermore, no axiomatic theory is used to specify the functional form of compaction function. The degrees/ problem of freedom can be overcome by expanding the evoked set with abstract alternatives specified only by their attributes. Based on axioms which specify whether

|                              | Linear Regression | Monotonic Regression |
|------------------------------|-------------------|----------------------|
| **Coefficients (t statistic)** |                   |                      |
| Quality - Raw                | 6.17 (9.16)       | .635                 |
| Normalized                   | .32               | .33                  |
| Personalness - Raw           | 3.86 (6.4)        | .495                 |
| Normalized                   | .20               | .26                  |
| Value - Raw                  | 5.69 (12.2)       | .494                 |
| Normalized                   | .30               | .26                  |
| Convenience - Raw            | 3.34 (6.0)        | .326                 |
| Normalized                   | .18               | .17                  |
| **Fit Measures**             |                   |                      |
| First Preference             | .48               | .52                  |
| Rank Order                   | .44               | .47                  |
| $R^2$                        | .27               | --                   |
| $\theta$                     | --                | .36                  |
| $\tau$                       | --                | .38                  |
| **Saved Data Fit**           |                   |                      |
| First Preference             | .36               | .40                  |
| Rank Order                   | .40               | .42                  |

Table 7:  Preference Regressions

a function is measurable on a data set (Tversky [62]) conjoint analysis

estimates individual compaction functions and derives importance weights by

having consumers rank order factorially generated sets of these abstract

alternatives. Although the measurement theory allows for polynomial com-

binations of nonlinear functions of the attributes most marketing applications

have been limited to additive or multiplicative combinations of these functions

(Green and Wind [19], Johnson [29]) or simple interactions (Green and Devita

[17]). In these cases the attributes have been instrumental (e.g., price,

brand name, package design) so that factorial combinations can be unambiguously

defined and design tradeoffs can be explicitly made with respect to controll-

able variables. Conjoint analysis was not applied to the HMO case since at the

early design stage interest is on psychological dimensions rather than instru-

mental variables. A study is now being conducted to link instrumental variables

such as HMO name, building, and waiting time to the perceived quality. For

an example of conjoint analysis applied to health and hospital choice, the

reader is referred to Wind and Spitz [70].

Since the conjoint axioms deal with measurement, they do not specify what

the functional form of the compaction function should be. On the other hand,

von Neumann-Morgenstern utility analysis draws on deductive theory to derive

unique functional forms from fundamental axioms and verifyable assumptions de-

scribing how consumers evaluate alternatives. These forms are important because

they allow direct measurement of risk aversion, importances, and interactions.

However, until now, utility theory has been used exclusively for prescriptive

decision making. Compaction adds the requirement that the theory allow stochastic

choice, i.e., that the probability of choosing a utility maximizing alternative is

not necessarily certain. To use utility theoretic results for descriptive choice

a theoretical construct of stochastic choice preference was defined ($a_1$ preferred to $a_2$ means the probability of choosing $a_1$ is greater than the probability of choosing $a_2$) and it was found necessary to augment the von Neumann-Morgenstern [67] axioms with a psychological choice axiom (Hauser [23]) which can be shown to be similar to simple scalability (Luce [40], Krantz [37], Tversky [62]). These additions enabled the establishment of an isomorphism between utility and compaction theory and thus the ability to apply many utility theoretic results including identification of functional forms and direct assessment to descriptive compaction theory. The full proofs are contained in Hauser [23].

Once the functional form is known measurement proceeds as follows. Rather than ranking factorially generated abstract alternatives, the consumer is asked to consider abstract alternatives two at a time. One alternative is completely specified, the other alternative leaves one characteristic (an attribute level or an indicator of uncertainty) unspecified. The consumer's task is to specify that characteristic so that he is indifferent between the two alternatives. If we assume that indifference means the consumer is equally likely to choose either abstract alternative, then the compaction values of the two alternatives are equal. Since the functional forms are known this provides one equation in the parameters. With sufficient indifference questions we can algebraically solve for the parameters. Calculations are based on the fact that $\underline{x}_{ij}'$ indifferent to $\underline{x}_{ij}''$ means $c(\underline{x}_{ij}', \underline{\lambda}_i) = c(\underline{x}_{ij}'', \underline{\lambda}_i)$.

As a preliminary test of this technique, compaction functions over the four performance measures describing health care delivery were assessed by a personal interview for a random sample of 80 members of the MIT student population. The compaction function was approximated with a relatively simple

functional form, which was separable, multiplicative, and constantly risk averse in each dimension. Mathematically, the function, with individual specific parameters, is stated here for four performance measures.

$$c_{ij} = \sum_m k_{im} u_{im}(x_{ijm}) + \sum_m \sum_{\ell > m} K_i \, k_{im} \, k_{i\ell} \, u_{im}(x_{ijm}) \, u_{i\ell}(x_{ij\ell})$$

$$+ \ldots + K_i^3 \, k_{i1} \, k_{i2} \, k_{i3} k_{i4} \, u_{i1}(x_{ij1}) \cdots u_{i4}(x_{ij4})$$

$$\text{with} \quad u_{im}(x_{ijm}) = [1-\exp(-r_{im} x_{ijm})]/[1-\exp(-r_{im} x_m^*)]$$

Where

$c_{ij}$     = individual i's scalar measure of goodness for alternative $a_j$, i.e., the value of the compaction function, $c_j(\underline{x}_{ij}, \underline{\lambda}_i)$ evaluated for individual i and alternative $a_j$.

$u_{im}(x_{ijm})$     = uni-attributed conditional "utility" scaling function. The form shown here is for constant risk aversion, and is scaled from $u_{im}(0) = 0$ to $u_{im}(x_m^*) = 1$. (If $r_{im} \to 0$, $u_{im}(\cdot)$ becomes the linear form used in statistical compaction.)

$x_{ijm}$     = the level of the $m^{th}$ performance measure as perceived by individual i for alternative $a_j$.

$x_m^*$     = the maximum value of the $m^{th}$ performance measure.

$r_{im}$     = individual i's risk aversion coefficient relative to the $m^{th}$ performance measure.

$k_{im}$     = individual i's importance coefficient for the $m^{th}$ performance measure.

$K_i$     = individual i's interaction coefficient relative to the four performance measures.

Before assessment, the multiplicative form was selected based on prior theory and in-depth interviews aided by an interactive utility assessment computer program developed by Sicherman [55]. In full scale assessment, independence questions were used to check the validity of the assumptions necessary for

the multiplicative form. It was found that these assumptions were correct
for 66% of the respondents. Since this is the first time these assumptions
have been tested on a consumer population this result is encouraging.
First we note that these assumptions are implicit in
the functional forms used in conjoint analysis. Second we note that there
are functional forms (Keeney [30], Fishburn [12], Farquhar [10]) which
relax these assumptions, but as yet it is not feasible to construct simple
consumer measurements for these forms. Based on the results of administering
the utility questionnaire (see "measurement" section of this methodology) the
parameters $\underline{\lambda}_i = \{k_{i1}, k_{i2}, k_{i3}, k_{i4}, r_{i1}, r_{i2}, r_{i3}, r_{i4}, K_i\}$ were calculated
and are shown in Table 8.

[Insert Table 8 here.]

An empirical result unexpected in utility theory is that the risk aversion
coefficients and importances are highly correlated. This indicates that the student
group is more concerned with risk for the more important performance measures.
As a first comparison against the statistical technique, the individual
specific perceptions, $\underline{x}_{ij}$, and preference parameters, $\underline{\lambda}_i$, were used to
calculate scalar measures of goodness, $c_{ij}$, for each alternative for each
individual. When compared against rank order preference these resulted in
a rank order fit of .474 and a first preference fit of .495. These are in
the same range as the fits of statistical procedure applied to the overall
sample. Because of the differences between the statistical and direct
compaction techniques especially in their relation to the rest of the
methodology, and because of the nonlinear relationship between the factor
scores and the directly measured performances, and because of the risk averse
scaling function, stronger, and more explicit comparison tests need to be
devised before importances can be compared. This is the subject of future

| | Importances ($k_{im}$) | | | | Risk Aversion Coefficients ($r_{im}$) | | | | Interaction Coeff. ($K_i$) |
|---|---|---|---|---|---|---|---|---|---|
| | quality | personal-ness | value | convenience | quality | personal-ness | value | convenience | |
| Median | .800 | .525 | .706 | .682 | .693 | .332 | .424 | .310 | -.992 |
| Inter-quartile interval | .650 - .900 | .360 - .726 | .492 - .800 | .525 - .800 | .424 - 1.10 | .000 - .693 | .203 - .549 | .000 - .693 | -.998 - -.939 |
| Inter-quartile range | .250 | .367 | .308 | .275 | .675 | .693 | .347 | .693 | .059 |

Table 8: Compaction by Utility Theoretic Procedures For MIT Students

research work (Hauser and Urban [25]).

In summary, compaction identifies how consumers use the performance measures to evaluate alternatives. The key idea is that the consumer compacts the performance measures into a scalar measure of goodness for each alternative which then compared across alternatives yields preference among the alternatives.

## Segmentation

Development of a single product or service may not be the best strategy to exploit a potential market. For example, some people may prefer high quality service and are willing to pay the price while others want adequate service at a low price. It may be that average service at an average price satisfies neither of these two groups. Thus, in the design of new products and services it is necessary to determine whether everyone has the same preferences relative to the reduced performance measures, or whether there exist segments of the population which have significantly different preferences. In the models of this methodology, this type of segmentation is represented by significantly different importance weights between segments, i.e., significantly different $\Lambda_s$'s.
This is a very attractive segmentation method since these parameters contain the key information used by the design team to determine managerial tradeoffs in the attributes of the new product or service. This type of segmentation has been conceptually proposed by Halley [22] and is called "benefit" segmentation. The proposed segmentation is in contrast to "life style" and "psychographics" as exemplified by Wells and Gruber [68]. The reader is referred to

Frank, Massy, and Wind [14] for a summary of the literature on segmentation. Although these methods may be useful in creating ads for established products we feel benefit segmentation is the best method for making attribute tradeoffs inherent in new product design. After benefit segments are defined, they can be described by a profile of demographic and psychometric identifiers. The relative importance of these descriptions in defining segment membership can be examined by discriminant analysis (Massy [44]).

In addition to the need for segmentation from a marketing strategy point of view, homogeneous groups are necessary for valid compaction if the preference regression technique is used. In the statistical estimation of compaction functions a fundamental paradox exists. It is theoretically sound to estimate importance weights only if the population is homogeneous with respect to these weights, but we do not have enough degrees of freedome for estimation unless we group individuals together. Therefore we must segment the market before statistical estimation.

The description of segments has been often highly judgmental. Techniques such as clustering (Green and Carmone [16]) and AID (Sonquist, Baker and Morgan [57]) have been used, but they are highly manipulative techniques and potentially misleading. See Doyle and Fenwick [ 9] for a discussion of AID dangers. We propose that these techniques can be used in the search for alternate segments, but that specific statistical tests are needed to determine the validity of the segmentation. We propose a statistical test that requires (1) the correlations to preference to be higher in each segment and (2) the importance coefficients to be significantly different across segments.

In the HMO case cluster analysis and AID analysis techniques were used to search for segments. AID was used with the measured intent to join the

MIT plan as the dependent variable and individual descriptors and demographics as the explanatory variables (Greer and Suuberg [21]). There were 690 observations and, therefore, somewhat less than the recommended AID sample of 1000. The most significant segmentation variables were measures of an individual's current pattern of care (e.g., currently MIT health department versus use of private doctor for physicals and continuing care). However, the evidence was not very strong. The AID analysis explained 24% of the variation while simulations based on random data explained 16%. Other attempts at segmentation were equally unconvincing. One attractive method of segmentation is to cluster the individual parameters of the utility function directly (see Table 8), but since the sample size for the utility study was only 80 students, clustering could not be utilized. This technique deserves future research effort.

Throughout the search for segmentation there was a common, but weak indication that the pattern of existing care - MIT versus private - was a possible variable for segmentation. The priors of the health department were that the segmentation of faculty, students, and staff was important since these were operational segments and it was felt that preferences varied across these groups.

Two alternative partitions were proposed and tested. The first was the prior of segmentation of faculty-students-staff while the second was based on the multivariate analysis and was private versus MIT as an existing care supplier. Table 9 presents the statistical regression across individuals and stumli of preferences against the perceived attributes. If a segmentation is real we require that (1) the regression fits to be significantly better statistically in each segment than in the overall regression and (2) the

importance coefficients to be significantly different across segments. An informal examination of Table 9 does not show much evidence of segmentation. The $R^2$ and F's are reported on the basis of the finding that monotonic regression had little effect on the coefficients (see Table 7). The $R^2$ and t's are significant but are not uniformly better within the segments. Only the faculty segment $R^2$ is improved substantially ($R^2 = .27$ overall versus .38 for faculty), but in this partition the $R^2$ for staff decreases substantially ($R^2 = .18$ versus .27 overall). Some differences in importances are evident between faculty and students with respect to personalness ($\lambda$ for faculty = 5.95 versus 1.69 for students). The $R^2$ in the second partition (private, MIT, or MIT and Private existing care patterns) are almost the same as the overall value.

[Insert Table 9 here.]

A formal statistical test has been described by Johnston [29] and Fisher [13] to test the significance of the difference of regressions. Assuming interval data and applying Johnston's test, none of the segments in the partitions are significantly different at the 10% level ($F_{195}^{10} = .77$ for faculty/student/staff partition and $F_{195}^{10} = .50$ for the private/MIT/Mixed pattern of care partition).

Thus, in the HMO data there is no statistical basis for benefit segmentation. This emphasizes the need for a formal statistical technique to test the segmentation identified by priors or the interpretive clustering or AID techniques.

| | $R^2$ | F(df) | Quality(t) | Personalness(t) | Value(t) | Convenience(t) |
|---|---|---|---|---|---|---|
| **Overall** | | | | | | |
| (n=210) | .27 | 80.0 (4,839) | 6.17 (9.16) | 3.86 (6.4) | 5.69 (12.2) | 3.34 (6.0) |
| | | | | | | |
| **Partition One** | | | | | | |
| Faculty (n=67) | .38 | 41.2 (4,263) | 6.04 (5.61) | 5.93 (5.61) | 5.95 (9.60) | 3.34 (3.5) |
| Student (n=80) | .31 | 36.3 (4,315) | 6.59 (6.10) | 1.69 (1.87) | 7.82 (8.26) | 4.38 (4.38) |
| Staff (n=63) | .19 | 14.3 (4,247) | 4.98 (3.60) | 5.15 (4.32) | 3.54 (3.48) | 2.08 (2.06) |
| | | | | | | |
| **Partition Two** | | | | | | |
| Private (n=88) | .28 | 34.0 (4,347) | 6.05 (5.66) | 4.60 (5.3) | 4.86 (7.2) | 4.53 (5.41) |
| MIT (n=109) | .29 | 44.0 (4,431) | 6.89 (6.91) | 3.54 (3.76) | 6.5 (3.76) | 2.79 (3.57) |
| Mixed (n=12) | .26 | 3.9 (4,43) | 2.59 (1.0) | 5.49 (2.0) | 4.14 (1.62) | -1.03 (.41) |

Table 9:  Segmentation

## Probability of Choice

Perception, compaction, and segmentation provide diagnostics necessary to design successful new products and services, but to evaluate potential designs we must focus on estimating how many consumers will choose the new product or service and how many of the potential consumers will select each of the various competing alternatives. The probability of choice module of the methodology explicitly models how the consumer compares the scalar measures of goodness computed in the compaction module. If we are far enough along in the design of a new product or service to have advertising copy and a sample product or pilot service available, the probability of choice model is estimated by observing how consumers choose among actual alternatives. Otherwise, as is the case early in the design process, actual new products or services are not yet available. In this case, scenarios are given to approximate choice as closely as is feasible. For example, on the first cycle through health services may be represented by concept statements, on the second cycle through we may have finished brochures and video tape testimonials, and on the third cycle a pilot program. Thus, when choice is possible we observe it, but when it is not possible we observe choice on proxy alternatives and augment the data with measures of intent.

Although the emphasis in early design of the HMO is on the discovery of perceptual dimensions, importances, and segments, a preliminary sales forecast is needed to see if the venture is viable. In the HMO case observed preference and intent measures are used to estimate actual choices.

Intent to enroll for MIT was measured on a 1-5 point scale (definitely yes, probably yes, might or might not, probably not, and definitely not). These intent measures were processed  on the assumption that all the definite

intent, half of the probable intent, and thiry percent of the might intent could be converted to choice at MIT. These are a little higher than those found for Gruber [20] for consumer brands (definite .755, probable .314, might .268), but we felt the MIT HMO marketing would result in more conversion and were consistent with the author's experience in test marketing [65]. The application of intent conversions resulted in an estimate that 23.3% of those aware of the HMO would enroll.

For consistency, the observed preference was converted to choice on the assumption that 80 percent of the first preference and 20 percent of the second preference would result in choice. These fractions are based on translations of preference to choice in simulated shopping studies of consumer products (Silk and Urban [56]). Based on these proportions it was found that 21.7% of those aware would enroll in the MIT HMO. Since this was consistent with the intent transformation value of 23.3%, a value of 23% was used in forecasting for the managerial decisions reported in a following section. In simulating the effects of a new design, the compaction coeffic- ients (see Table 7) were used along with new perceptual attribute values to predict individual preference rank order. These were converted to choice by assuming 80 percent of first preference and 20 percent of second preference would choose MIT if they were aware.

The choice can be considered as the outcomes of Bernoulli probabilities, $(p_s(a_j|c_{i1}, c_{i2}, \ldots, c_{ij}, \ldots, c_{iJ}))$. These estimate each individual's selection probabilities within a segment (s) for each alternative $(a_j)$ conditioned on the scalar goodness measures for each alternative $(c_{i1}, c_{i2}, \ldots, c_{ij}, \ldots, c_{iJ})$. In cases where repetitive choice decisions are made by a consumer, separate trial and repeat choice parameters would be estimated based on the goodness measure before use and after use of the new product or service.

One approach is to use the multinomial logit model (McFadden [42]). This model postulates that there is a true "utility", $u_{ij}$, which completely describes a consumer's choice process (i.e., a consumer chooses the alternative, $a_\ell$, with the largest $u_{i\ell}$), but we can only observe part of $u_{ij}$. In fact, the "true" utility is equal to an observable part, $c_{ij}$, plus a random error. Deductive reasoning from distributional assumptions on the error term yields the multinomial logit model:

$$p(a_j | c_{i1}, c_{i2}, \cdots c_{iJ}) = e^{\beta c_{ij}} / \sum_{\ell=1}^{J} e^{\beta c_{i\ell}}$$

There are two ways to use this model in the methodology. The first is to substitute the compaction values directly into the logit model and estimate the parameter $\beta$ by maximum likelihood. The second method is to specify the compaction function as a linear or non-linear function of attributes and estimate parameters ($\lambda_i$) by maximum likelihood. In this case $\beta$ in the logit model is incorporated in the $\lambda_i$. If $c_s(\underline{x}_{ij}, \underline{\lambda}_i)$ is linear in the parameters, a number of programs are readily available (McFadden [41]).

The $\beta$-logit model has been calibrated and tested on data collected prior to national introduction of a new aerosol deodorant. The scalar measures of goodness were constant sum paired comparison scale values and the choice was the respondent's last purchase (Silk and Urban [56]). In this example, $\beta$ was 2.09 with a t statistic of 10 (df = 278). The goodness of fit of the model was evaluated based on a new set of entropy tests for a probability of choice model.

## Testing probability of choice models

Whether done on proxy choice and intent measures, or done on actual choice with the multinomial logit, the output of the probability of choice

module is a predicted probability for each individual consumer for each
choice alternative. In compaction we tested alternative models based on
their ability to recover either first preference or rank order preference;
in probability of choice we would like another measure. We would like to
discriminate between a model which assigns .9999 to the chosen alternative
over a model which assigns .5001 to an alternative. Traditional measures
break down for this test because our predicted measures, probabilties, are
scaled between 0 and 1 while our observed measures, choices, are nominally
scaled (0 or 1). In fact, for this type of data it can be shown (Hauser
[24]) that the expected value of tests such as least squares are optimized
for values other than true probabilities. Instead we propose a test
described in Hauser [24] based on honest reward theory (Raiffa [50]) and
information theory (Gallagher [15]). This test begins with a naturally
occuring measure of uncertainty, entropy, and computes the percent of that
uncertainty that is explained by the probability model.

Specifically, if our null hypothesis is that all individuals have the
same probability of choosing $a_j$ then the entropy, $H_o(A)$, relative to $N_o$ is
given by

$$H_o(A) = - \sum_{j=1}^{J} m_j^o \log m_j^o$$

where $m_j^o$ is the null probability of choosing $a_j$. The observed information,
$I(A;X)$, given by the model relative to the null hypothesis is given by:

$$I(A;X) = \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{J} \delta_{ij} \log (P_{ij}/m_j)$$

where n = the total number of consumers, $P_{ij}$ are the probabilities predicted
by our model, and $\delta_{ij}$ = 1 if i chooses j and 0 otherwise. The percent of
uncertainty, $U^2$, explained by the model relative to the null case is simply

$I(A;X)/H_o(A)$ and the residual uncertainty is simply $1 - U^2$.

In the deodorant case, this information test was applied. The null hypothesis, $N_o$, was that the probability of choice was proportional to the empirical share of people who chose each alternative. The $U^2$ value was 74% ($H_o(A) = 2.23$ and $I(A;X) = 1.64$). See Hauser [24] and Silk and Urban [56] for discussion and application of this test. This indicates substantial contribution of the model to reduction of uncertainty and along with the t statistic for $\beta$ of 10, suggests highly sifnificant results. Probability of choice predicts individual choice for products or services in an individual's evoked set; the next step aggregates the individual choice probabilities.


## Aggregation

The final step in the consumer response process model is aggregation. It combines the individual choice probabilities to produce numerical estimates of the total share of choices and number of people choosing each alternative. If relevant population segments were identified in abstraction, aggregation explicitly uses them to extrapolate from the sample population to the target population. In addition to expected choices, the variances on the group choice is also useful in considering the risk associated with the new venture.

In most applications, individual choice probabilities are roughly independent and thus the Central Limit Theorem can be used to calculate the joint probability distribution of the market shares. The grand means are determined by averaging the individual results obtained by direct substitution of the individual goodness measures in the probability of choice models ($P_{ij}$). The variance is based on the sum of the individual variances

$(P_{ij}(1-P_{ij}))$. In some cases it is convenient to represent a distribution across the goodness measures in which case one must integrate rather than sum to find the mean and variance (see Koppelman [36]). Explicitly the mean market share for alternative $a_j$ is given by:

$$\overline{ms}_j = \frac{1}{n} \sum_{i=1}^{n} P_{ij}$$

The variance is given by:

$$var(ms_j) = \frac{1}{n} \sum_{i=1}^{n} P_{ij}(1-P_{ij}).$$

The final step in aggregation reflects the fact that all the measurement and estimation is based on consumers who evoked the alternative by past experience or who were made aware of the alternative in the interview. The adjustment is to multiply the aggregate share of choices for an alternative times the estimated level of evoking and awareness that the new product or service is expected to achieve based on the predicted level of marketing effort or the evoking rates for competing alternatives.

In the HMO case, the predicted probabilities were discrete so the summation technique was used. Evoking and awareness were estimated judgmentally based on a planned marketing effort. The empirical results are discussed in the case application.

This completes the analytic discussion of the methodology. The primary purpose of the methodology was to aid in the design of new products and services. The next section illustrates this use by discussing the managerial implications of the HMO case.

## 5. <u>Case Application</u>

The proposed methodology for modeling consumer response to innovation has undergone an application to the problem of converting the MIT health department into an HMO. Some of the empirical findings have been cited earlier. This section will concentrate on the managerial use of the model in evaluation and refinement of the HMO design.

The model estimation was based on consumer interviews of 447 faculty, students, and staff. Of these 367 were prospective members of the HMO and 80 were members of a pilot HMO begun a year earlier.

First perceptual maps were derived by a factor analysis of the ratings of existing and new health plan descriptions. Figure 5 gives the overall average factor scores (see Table 1 for a concept statement, Table 2 for the rating scales, and Table 6 for factor loadings).

[Insert Figure 5 here.]

First, it should be noted that the average perceptions of the existing care system are better than those of the MIT HMO concept based on the measures from the prospective members on all dimensions except convenience. Next, note that the MIT HMO concept is perceived better than the Harvard Community Health Plan (HCHP) on all dimensions except quality. The examination of the factor score coefficients and raw ratings (Table 5) showed the lower quality rating for the MIT HMO was based almost entirely on a low score for MIT on hospital quality. This was because the MIT plan hospitals were smaller and not well regarded (i.e. Cambridge City Hospital) as compared to the prestigious Boston hospitals. The Harvard Community Plan was higher on quality based on its hospital ratings (Beth Israel, Peter Bent Brigham, and Childrens Hospital).

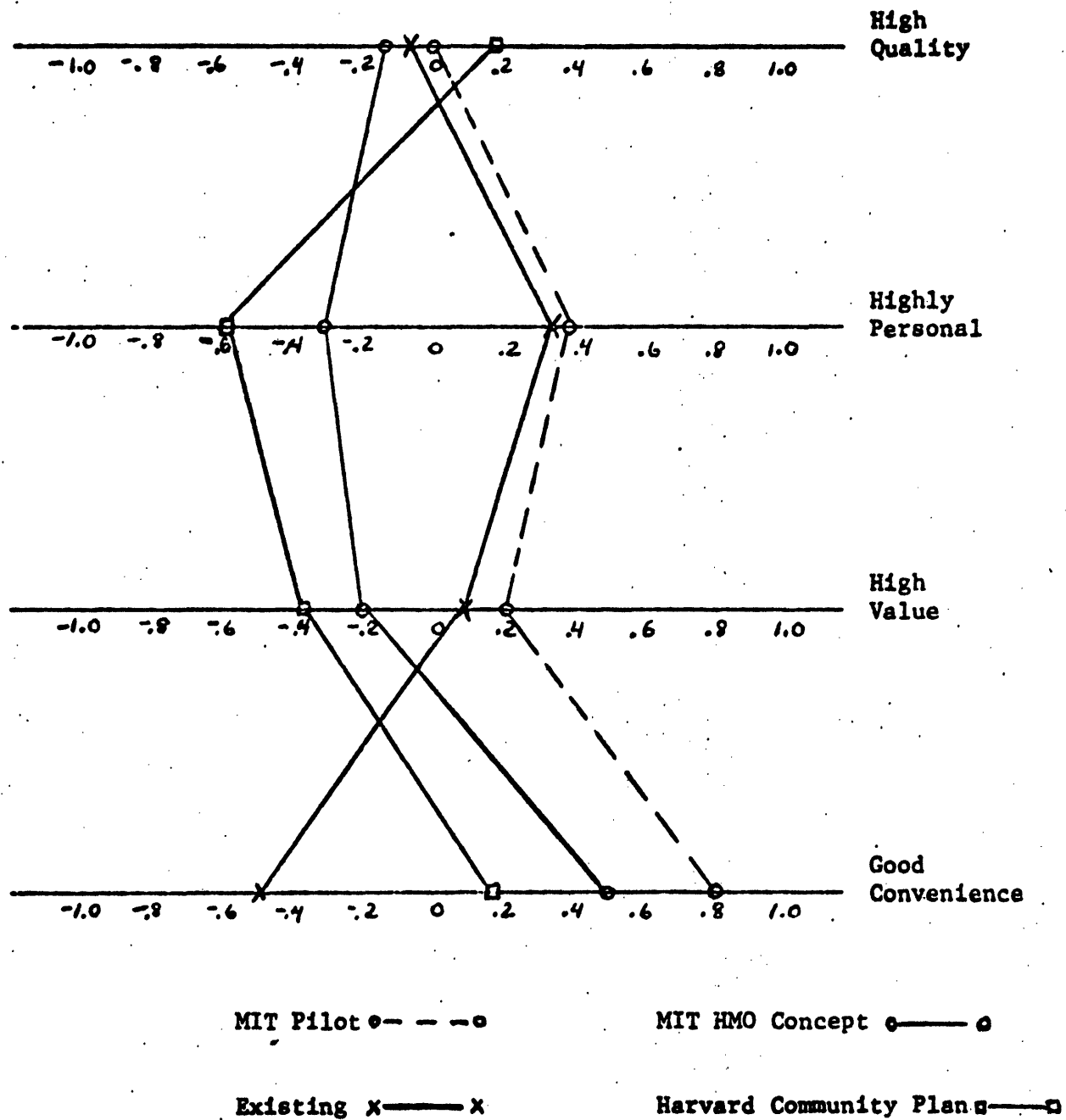When comparing existing care perceptions with the perceptions of

**High Quality**

-1.0   -.8   -.6   -.4   -.2      .2   .4   .6   .8   1.0

**Highly Personal**

-1.0   -.8   -.6   -.4   -.2   0   .2   .4   .6   .8   1.0

**High Value**

-1.0   -.8   -.6   -.4   -.2   0   .2   .4   .6   .8   1.0

**Good Convenience**

-1.0   -.8   -.6   -.4   -.2   0   .2   .4   .6   .8   1.0

MIT Pilot ●— — —●          MIT HMO Concept ●———— ●

Existing X————X          Harvard Community Plan □————□

Figure 5:  Overall Average Factor Scores

those in the MIT pilot plan, a substantially different picture emerges. The MIT pilot plan exceeds the average existing care perceptions on all dimensions. The stated intent to re-enroll of over 90% supported the notion that the plan was very effective as seen by those in the pilot plan. It is clear that perceptions based on actual pilot plan performance of the HMO are much better than perceptions of the plan based on the concept statement. Although this could be post purchase rationalization, it is more likely that this is a case of a good product where few people perceive it as such until using it.

The perceptual maps indicated two major managerial findings. First, "quality" of the plan was low and probably would be improved by better hospital affiliation and second, if the HMO was to be successful it would have to develop an aggressive campaign to communicate actual plan performance to perspective members.

Analysis of the preferences indicated that overall the rank order of importance of the four relevant attributes was quality, value, personalness, and convenience.

The choice model and empirical observations were used to forecast new enrollment and re-enrollment for the next year. Table 10 gives the forecast based on the estimated probability of enrollment given that the respondent was aware at the level presented by the survey and based upon an estimate of how many potential users would be aware at this level. Since the choice probabilities are different in each group, the new enrollment is based on each component segment. Re-enrollment is based on the empirically estimated probability of re-enrollment (92.5%) and estimates of migration

out of the MIT community. The total enrollment forecast of 3600 families was just financially sufficient to maintain the HMO. Considering the inherent risk involved in any new service ventrue, the decision to expand the existing pilot HMO could not be supported based on the initial design.

However, the existing pilot was not the best design. The use of aggressive communication to close the gap of perceptions and performance and the change of hospital associations were identified as methods of improvement. The model was used to simulate the effects of these design changes by assuming (1) the communication campaign and the association with prestigious hospitals could move the perceptions one half of the distance from concept to actual pilot performance on quality, personalness, and value (see Figure 5) and, (2) 85% awareness would be created rather than 70%. The consumer response model forecast an enrollment of 5400 families based on estimates of importance weights and a linear compaction function produced by the regression of factor scores against preference (see Table 7) and the conversion of rank order preference to choice described in the compaction section of this paper. Other simulations were done based on assumptions of competition from the Harvard Community Health Plan (HCHP) which indicated

- [INSERT TABLE 10 HERE] -

the enrollment could drop to 2400 with HCHP offered at MIT and no improved design. The most likely forecast was based on HCHP being offered with an improved MIT plan and was for 4950 family enrollments. This was not suffi- cient to make a positive recommendation for full implementation and a new building even based on the response of consumers to the revised communication and design strategy. The predictions did substantiate the demand for more

## CASE I - EXISTING DESIGN

**New Enrollment:**

| Groups | Number not now in pilot HMO | | Enrollment if aware | | Estimated Awareness | | Estimated Enrollment |
|---|---|---|---|---|---|---|---|
| Students | 8000 | x | 33% | x | 70% | = | 1,848 |
| Faculty | 3800 | x | 15% | x | 70% | = | 399 |
| Staff | 3400 | x | 22% | x | 70% | = | 523 |
| Overall | 17,200 | | 23% | | 70% | | 2,770 |

**Re-Enrollment:**

| Existing HMO Subscribers | | Repeat Rate | | Estimated to Remain at MIT | | |
|---|---|---|---|---|---|---|
| 1067 | x | 92.5% | x | 86.3% | = | 852 |

**Total Enrollment** = 3622


## CASE II - IMPROVED DESIGN

**New Enrollment:**

| Groups | Number not now in pilot program | | Enrollment if aware | | Estimated Awareness | | Estimated Enrollment Number |
|---|---|---|---|---|---|---|---|
| Students | 8000 | x | 42% | x | 85% | x | 2,856 |
| Faculty | 3800 | x | 25% | x | 85% | x | 808 |
| Staff | 3400 | x | 30% | x | 85% | x | 867 |
| Overall | 17,200 | | 31% | | 85% | | 4,531 |

**Re-Enrollment:**

| Existing HMO Subscribers | | Repeat Rate | | Estimated to Remain at MIT | | |
|---|---|---|---|---|---|---|
| 1067 | x | 95% | x | 86.3% | = | 874 |

**Total Enrollment** = 5,405


Table 10:  Forecast of MIT Enrollment

HMO services and supported a plan to expand within the existing physical facilities to 3000 faculty and staff families. During this expansion the hospital affiliation was shifted from Cambridge City Hospital to Mount Auburn Hospital with referral to Massachusetts General Hospital. MIT is now expanding its HMO to meet the indicated need and will consider marketing its HMO to students as facilities become available.

Based upon this initial application it appears that the consumer response model is relevant to the management of innovation and can be useful in improving designs of new products or services, forecasting the acceptance of innovations, and reducing the risk of failure.

## 6. Summary and Final Comments

This paper proposes and presents evidence for a normative methodology to elicit and guide creativity in the design of innovative products and services. The strength of the methodology is that it effectively integrates state of the art analysis techniques from the fields of psychometrics, utility theory, and stochastic choice theory in a model based analysis process oriented toward the needs and desires of managers and staff responsible for innovation. Its primary use is to enhance early creative identification and design of high potential products and services by providing important diagnostics which describe consumers' perceptions of the alternatives and consumers' preferences relative to measures of these perceptions. It also identifies managerially relevant segments based on homogeneity of preference and gives numerical indications of consumer response within each segment. Diagnostics are produced for design insight, but the methodology also simulates and evaluates quantitative and qualitative

design changes in the characteristics of the alternatives or in the imple-

mentation strategy. Thus managers can readily test and improve the intuition

they develop based on previous experience and on the previous outputs of

the methodology.

The new techniques described in the methodology will be tested in

future work. Attention will be directed at confirming the validity of

(1) the new measurement instruments developed to allow mass direct assessment

of consumers' utility functions, (2) the use of psychometrics to get a

complete and parsimonious set of performance measures for utility assessment,

(3) the criteria and statistical test of segmentation based on preference

parameters, (4) the stochastic interpretation and choice axiomization which

allows an isomorphism between utility theoretic results and compaction,

and (5) the information theoretic testing techniques for assessing the

accuracy of predicted individual probabilities.

In addition to testing the new techniques proposed in this paper,

comparative research is needed on the alternative methods of compaction.

The technique of expectancy value, preference regression, conjoint analysis,

and utility theory should be analyzed to determine theoretically and

empirically when each technique is most appropriate. Some work is underway

(Hauser and Urban [ 25 ]). Utility theory needs new measurement methods

that are less demanding on the consumer and can be executed in the relatively

short time available in a personal interview. In addition, the empirical

results from applying utility theory in the HMO case indicates error in fit

and predictions, yet utility theory has no structure to deal with measurement

error. This is a high priority area for research. The multinomial logit

model proposed for the modeling of probability of choice is basically a zero order model. The probability of purchases does not depend in past purchasing except through preference values. In many situations switching and purchase loyalty are phenomena to consider. (Massy, Morrison, Montgomery [ 45 ]). Research is needed in the formulation of higher order stochastic models that include preference and switching phenomena. A final research area is in the complementary use of conjoint analysis and perceptual mapping. Research is needed to develop procedures so conjoint analysis can be used to link the instrumented variables of the product to the perceptual performance dimension.

The methodology has been applied in the repositioning of the master of science program at MIT's Sloan School of Management, to the design of financial service packages (banking and brokerage), and the positioning of new frequently purchased consumer products (e.g. antacids, personal care products, and pain relievers). In these ten applications managerial reception has been enthusiastic and several high potential products have been designed. Applications are now planned to develop consumer acceptance of squid as a protein source and to design both suburban and intercity passenger transportation systems. Although the initial experience is encouraging future application, testing and research will be needed to fully document and develop the impact of applying the proposed methodology to the design of new products and services.

## Bibliography

1. Allaire, Yvan, "The Measurement of Heterogeneous Semantic, Perceptual, and Preference Structures", (Unpublished Ph.D. Dissertation, MIT, 1973).

2. Ashton, Winifred D., The Logit Transformation, (London: Griffin Co., 1972).

3. Bass, Frank M. and W. Wayne Talarzyk, "An Attitude Model for the Study of Brand Preference", Journal of Marketing Research, 9, (February 1972), pp. 93-96.

4. Ben-Akiva, M.E., Structure of Passenger Travel Demand Model, MIT, Civil Engineering, Ph.D. Thesis, 1973.

5. Booz, Allen, and Hamilton, Management of New Products, 1971.

6. Carroll, J.D., "Individual Differences and Multidimensional Scaling", in R.N. Shepard, A.K. Romney, and S. Nerlove, eds., Multidimensional Scaling: Theory and Application in the Behavioral Sciences, (New York: Academic Press, 1972), pp. 105-157.

7. Carroll, J.D. and J.J. Chang, "Analysis of Individual Difference in Multidimensional Scaling Via N-Way Generalization of 'Ekart-Young' Decomposition", Psychometrika, Vol. 35 (September 1970), pp. 283-319.

8. Carroll, J.D. and J.J. Chang, "Relating Preference Data to Multidimensional Scaling Solutions via a Generalization of Coombs' Unfolding Model", Bell Telephone Laboratories, Murray Hill, New Jersey, 1967 (mimeographed).

9. Doyle, P. and I. Fenwick, "The Pitfalls of AID Analysis", Journal of Marketing Research, Vol. XII (November 1975), pp. 408-413.

10. Farquhar, Peter H., "A Fractional Hypercube Decomposition Theorem for Multiattribute Utility Functions", Operations Research, Vol. 23, No. 5, Sept.-Oct. 1975, pp. 941-967.

11. Fishbein, M., "Attitudes and the Prediction of Behavior", in M. Fishbein, ed., Readings in Attitude Theory and Measurement, (New York: John Wiley and Sons, 1967).

12. Fishburn, Peter, "von Neumann-Morgenstern Utility Functions on Two Attributes", Operations Research, Vol. 22, No. 1, January 1974, pp. 35-45.

13. Fisher, Franklin M., "Tests of Equality Between Sets of Coefficients in Two Linear Regressions: An Expository Note", Econometrics, Vol. 38, No. 2 (March 1970).

14. Frank, R.E., W.F. Massy and Y. Wind, Market Segmentation, (Englewood Cliffs: New Jersey, Prentice Hall, 1972).

15. Gallagher, Robert, Information Theory and Reliable Communication, (New York: John Wiley and Sons, Inc., 1968).

16. Green, Paul E. and Frank J. Carmone, Multidimensional Scaling and Related Techniques in Marketing Analysis, (Boston, MA: Allyn and Bacon, Inc., 1970).

17. Green, Paul E. and Michael T. Devita, "An Interaction Model of Consumer Utility", Journal of Consumer Research, Vol. 2, Sept. 1975, pp. 146-153.

18. Green, Paul E. and Vithala R. Rao, Applied Multidimensional Scaling, (New York: Holt, Rinehart, and Winston, 1972), p. 125.

19. Green, Paul E. and Yoram Wind, Multiattribute Decisions in Marketing, (Hinsdale, Ill.: The Dryden Press, 1973).

20. Gruber, Alin, "Purchase Intent and Purchase Probability", Journal of Advertising Research, Vol. 10 (February 1976), pp. 23-28.

21. Greer, Peter S. and Eric M. Suuberg, "Segmentation Issues in Marketing and HMO", MIT, Sloan School of Management, S.M. Thesis, 1975.

22. Halley, R.I., "Benefit Segmentation: A Decision-Oriented Research Tool", Journal of Marketing, Vol. 32 (July 1968), pp. 30-35.

23. Hauser, John R., A Normative Methodology for Modeling Consumer Response to Innovation: Issues, Models, Theory, and Use, Unpublished Sc.D. Thesis, Operations Research Center, MIT, June 1975.

24. Hauser, John R., "The Information Measure Provides a Mathematically Rigorous and Intuitively Appealing Predictive Test for Probabilistic Travel Demand Models", Working Paper, Transportation Center, Northwestern University, Evanston, Illinois, March 1976.

25. Hauser, John R. and Glen L. Urban, "Direct Assessment of Consumer Utility Functions: von Neumann-Morgenstern Utility Theory Applied to Marketing", Working Paper #843-76, Sloan School of Management, Massachusetts Institute of Technology, 1976.

26. Hertz, David B., "Risk Analysis in Capital Investments", Harvard Business Review, (January-February, 1974), pp. 95-106.

27. Johnson, Richard M., "Tradeoff Analysis of Consumer Values", Journal of Marketing Research, Vol. XI (May 1974), pp. 121-27.

28. Johnson, Richard M., "Monotonic Regression", unpublished paper, (Market Facts, Inc.)

29. Johnston, J., Econometric Methods, (New York: McGraw-Hill, 1960), pp. 136-138.

30.   Keeney, Ralph L., "Multiplicative Utility Functions", <u>Operations Research</u>, Vol. 22, No. 1, January 1974, pp. 22-33.

31.   Keeney, Ralph L., "A Decision Analysis with Multiple Objectives: The Mexico City Airport", <u>Bell. J. Economics and Management Science</u>, Vol. 4, 1973, pp. 101-117.

32.   Keeney, Ralph L., "Utility Functions for Multiattributed Consequences", <u>Management Science</u>, Vol. 18, 1972, pp. 276-287.

33.   Keeney, Ralph L., "Utility Independence and Preference for Multiattributed Consequences", <u>Operations Research</u>, Vol. 19, 1971, pp. 875-893.

34.   Kelly, G.A., <u>The Psychology of Personal Constructs</u>, Vol. 1, (New York: Norton Wiley, 1955).

35.   Khlar, David, "A Monte Carlo Investigation of the Statistical Significance of Krusdal's Non-Metric Scaling Procedure", <u>Psychometrika</u>, Vol. 34 (September 1969), pp. 319-330.

36.   Koppleman, Frank, "Prediction of Travel Behavior with Disaggregate Choice Models", Unpublished Ph.D. Thesis, MIT Civil Engineering, June 1975.

37.   Krantz, D.H., <u>The Scaling of Small and Large Color Differences</u> (Ph.D. Thesis, University of Pennsylvania), Ann Arbor, Michigan, University Microfilms, 1964. No. 65-5777. As referenced by Tversky [63].

38.   Kruskal, J.B., "Multidimensional Scaling By Optimizing Goodness Of Fit to a Non-Metric Hypothesis", <u>Psychometrika</u>, Vol. 24, 1964, pp. 1-27.

39.   Little, John D.C., "Models and Managers: The Concept of a Decision Calculus", <u>Management Science</u>, May 1970, pp. 466-485.

40.   Luce, R. Duncan, <u>Individual Choice Behavior</u>, (New York: John Wiley and Sons, Inc., 1959).

41.   McFadden, D., "Quanta Choice Analysis: A Survey", NSF-NBER Conference on Decision Roles and Uncertainty, March 1974.

42.   McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior", in Paul Zaremblea, ed., <u>Frontiers in Econometrics</u>, (New York: Academic Press, 1970), pp. 105-142.

43.   Manski, Charles, "The Structure of Random Utility Models", Working Paper, Carnegie-Mellon University, Pittsburgh, PA, February 1975.

44.   Massy, W.F., "Discriminant Analysis of Audience Characteristics", <u>Journal of Advertising Research</u>, Vol. V, (March 1965), pp. 39-48.

45.   Massy, W.F., D.B. Montgomery and D.G. Morrison, <u>Stochastic Models of Buying Behavior</u>, (Cambridge, MA: MIT Press, 1970).

46. Pessemier, Edgar A., "STRATOP: A Model for Designing Effective Product and Communications Strategies", Working Paper No. 470, Purdue University, West Lafayette, Indiana, August 1974.

47. Quarmby, D.A., "Choice of Travel Modes for the Journey to Work", _Journal of Transportation Economics and Policy_, Vol. 1, No. 3, September 1967, pp. 273-314.

48. Raiffa, Howard, _Decision Analysis: Introducing Lectures on Choices Under Uncertainty_, (Reading, MA: Addison-Wesley, 1970).

49. Raiffa, Howard, "Assessments of Probabilities", Unpublished manuscript, January 1969.

50. Raiffa, Howard and Robert Schlaifer, _Applied Statistical Decision Theory_, (Boston, MA: Harvard University, 1961).

51. Richard, Scott F., "Risk Aversion, Utility Independence and Separable Utility Functions", _Management Science_, Vol. 22, No. 1, September 1975, pp. 12-21.

52. Rosenberg, M.J., "Cognitive Structure and Attitude Effect", _Journal of Abnormal and Social Psychology_, Vol. 53, 1956, pp. 367-372.

53. Ryan, Michael J. and E.H. Bonfield, "The Fishbein Extended Model and Consumer Behavior", _Consumer Research_, Vol. 2, No. 2 (September 1975) pp. 118-136.

54. Sheth, Jagdish N. and W. Wayne Talarzyk, "Perceived Instrumentality and Value Importance as Determinants of Attitudes", _Journal of Marketing Research_, 9, (February 1972), pp. 6-9.

55. Sicherman, Alan, "An Interactive Computer Program for Assessing and Using Multi-attributed Utility Functions", Unpublished Master's Thesis, MIT, Operations Research Center, February 1975.

56. Silk, A.J. and G.L. Urban, "Pretest Market Evaluation of New Packaged Goods: A Model and Measurement Methodology", Working Paper, Sloan School of Management, MIT, January 1976.

57. Sonquist, John, E. Baker and J. Morgan, "Searching for Structure", (Alias AID III), Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, Michigan, 1971.

58. Srinivasan, V. and Allan Shocker, "Linear Programming Techniques for Multidimensional Analysis of Preferences", _Psychometrika_, Vol. 38, (September 1973), pp. 337-370.

59. Stefflre, Volney, "Market Structure Studies: New Products for Old Markets and New Markets (Foreign) for Old Products", in Frank Bass, Charles King, and Edgar Pessemier, eds., _Applications of the Science in Marketing Management_, (New York: John Wiley and Sons, 1968), pp. 251-268.

60. Torgenson, Warren S., _Methods of Scaling_, (New York: John Wiley and Sons, 1958), pp. 105-112.

61. Tucker, L.R. and S. Messick, "An Individual Differences Model for Multidimensional Scaling", _Psychometrika_, Vol. 28, 1963, pp. 333-367.

62. Tversky, Amos, "A General Theory of Polynomial Conjoint Measurement", _Journal of Mathematical Psychology_, Vol. 4 (1967), pp. 1-20.

63. Tversky, Amos, "Elimination By Aspects: A Theory of Choice", _Psychological Review_, Vol. 79, No. 4, 1972, pp. 281, 299.

64. Urban, Glen L., "PERCEPTOR: A Model for Product Positioning", _Management Science_, VIII, April 1975, pp. 858-71.

65. Urban, Glen L., "SPRINTER MOD III: A Model for the Analysis of New Frequently Purchased Consumer Products", _Operations Research_, Vol. 18, Sept.-Oct., 1970, pp. 805-854.

66. Utterback, James, "Innovation in Industry and the Diffusion of Technology", _Science_, Vol. 183, pp. 620-626.

67. von Neumann, J. and O. Morgenstern, _The Theory of Games and Economic Behavior_, 2nd edition, (Princeton, New Jersey: Princeton University Press, 1947).

68. Wells, W.D. and G. Gruber, "Life Cycle Concept in Marketing Research", _Journal of Marketing Research_, Vol. III, November 1966, pp. 355-63.

69. Wilkie, William L. and Edgar A. Pessemier, "Issues in Marketings Use of Multi-Attribute Attitude Models", _Journal of Marketing Research_, Vol. 10, November 1973, pp. 428-41.

70. Wind, Yoram and Lawrence K. Spitz, "An Analytical Approach to Marketing Decisions in Health Care Organizations", _Operations Research_, Vol. 25, (forthcoming Sept.-Oct. 1976).

71. Young, F.W. and W.S. Torgerson, "TORSCA, A Fortran IV Program for Shepard-Kruskal Multidimensional Scaling Analysis", _Behavioral Science_, Vol. 12, 1967, p. 498.