Mathl Comput. Modelling, Vol. 10, No. 12, pp. 929–949, 1988
Printed in Great Britain.

0895-7177/88 $3.00 + 0.00

ISSUES AND ADVANCES IN PRODUCT POSITIONING MODELS IN MARKETING RESEARCH

D. H. GENSCH
School of Business, University of Wisconsin—Milwaukee, Milwaukee, WI 53201, U.S.A.

R. G. JAVALGI
School of Business, Cleveland State University, Cleveland, OH 44115, U.S.A.

(Received August 1987; accepted for publication May 1988)

Communicated by X. J. R. Avula

Abstract—Mathematical models used for product positioning are proliferating at a rapid rate. To provide a structure to this research, we have categorized the models/algorithms into three basic approaches: utility functions, maps and trees. In this article we identify and discuss the recent advances related to these approaches, discuss data requirements and indicate issues that influence that choice of the particular product positioning approach best-suited for various situations.

1. INTRODUCTION

Mathematical models for determining perceptual structure, which typically takes the form of maps, utility functions and/or hierarchical structures (trees), have dominated the current research on product positioning [1–5]. These types of models help practitioners to assess market opportunities by enabling them to discern the structure of the competition amongst products, which is indicated by their relative positions in trees or maps, as well as the relative preference for various attributes or levels of the attributes defining a product alternative. Furthermore, an understanding of perceptual structure methods can assist practitioners in identifying the "best" locations for existing or new products, and/or function as a guide in how to best alter existing products which is essential to advertising strategy and other components of the marketing mix.

The present paper's aim is to provide an informative discussion of important decision issues and advances germane to product positioning methods. This discussion is particularly important to analysts who often make seemingly innocuous decisions ranging from the type of data input to the choice of appropriate methods that can yield perceptual maps/spaces. While economists [6] use a space spanned by physical attributes, marketers [7] and psychometricians [8] have generally resorted to a perceptual space of reduced dimensionality. Perceptual mapping techniques include a large family of methods which can be used to create perceptual configurations of product markets. Analysis of such representations may be based on measures of overall similarity/dissimilarity, perception, preferences and the importance of the set of attributes that are embodied in the products. Customers or segments may also be represented in such a configuration by the location of their "most preferred" combination of attribute levels—termed their ideal point.

The decision issues and advances to be discussed in this article are organized according to the following broad categories:

I. Basic product positioning techniques

- Reduced-space compositional techniques such as: principal component analysis, factor analysis and discriminant analysis.
- Mapping techniques based on multidimensional scaling methods.
- Utility-based procedures that include compositional methods (e.g. expectancy value methods, regression, logit etc.) and decompositional methods (e.g. conjoint analysis).
- Hierarchial (nonoverlapping) and overlapping clustering methods.

II. Decision issues related to product positioning models
- Selection of number of dimensions based on: (i) *a priori* knowledge of salient dimensions on which choice is made; and (ii) number of attributes influencing choice and also homogeneity of the population.
- Determination of the dimensions which are not *a priori* known.
- Levels of aggregation: (i) issues (problems) related to individual level modeling; (ii) issues related to aggregate level analysis; and (iii) segmented modeling analysis.
- Interpretation of the diagnostic information generated from the output.

III. Factors influencing the type of models to use
- Knowledge of the problem area.
- Type of data available.
- Purpose of the study.

IV. Reliability and validity testing

2. BASIC STRUCTURE OF POSITIONING TECHNIQUES

Every product or service offered by a marketing organization has *attributes*, *characteristics* or *features* that are perceived by the consumer. Perceptual maps are based on a perception of the products (brands), which have attributes/features embodied within them. Perceptual maps are generated from the images consumers hold about different products. Methods for tapping these images may be categorized as:

(i) Reduced-space methods that utilize product attribute rating data to produce perceptual maps.
(ii) Utility-based procedures that include compositional and decompositional methods.
(iii) Procedures that ask samples of consumers to compare products (one pair at a time) and make judgments about the similarity (perception)/preference of those products. While making such judgments, consumers are assumed to rely on underlying dimensions. These dimensions and their underlying structures are inferred by a technique commonly known as *multidimensional scaling*.
(iv) Another approach that is distinct from the above two methods is known as *hierarchical clustering*.

The ensuing discussion focuses on these positioning methods.

Perceptual Mapping Based on Reduced-space Compositional Techniques for Product Positioning Using Multiattribute Rating Data

Perceptual mapping has received much attention in the current literature [1–4]. Perceptual maps define the dimensions and gaps where products (brands)† do not exist. Perceptual mapping analysis, based on reduced-spaces approaches that utilize brand attribute rating data, have become important and frequently used mathematical tools to analyze product positioning. The primary compositional-based techniques that have appeared in the literature and are widely used are: principal component analysis (PCA) [2–4, 10], factor analysis (FA) [2–4, 10], and (multiple) discriminant analysis (MDA) [11]. The use of these methods in marketing problems is widespread. Our presentation of these techniques is succinct. A more detailed discussion of these techniques may be found elsewhere [12–14].

Principle component analysis (PCA)

PCA tends to orient a space to dimensions that have high variance both within and between objects. The purpose of PCA is to determine factors (i.e. principal components) in order to explicate as much of the total variation in the data as possible, with as few factors as possible.

†In our study we often use the terms "products" and "brands" interchangeably. This terminology is adapted because it is generally not known *a priori* whether the entries that constitute the set of alternatives are from a "conventional" product class or are drawn from different classes (see Ref. [9]).

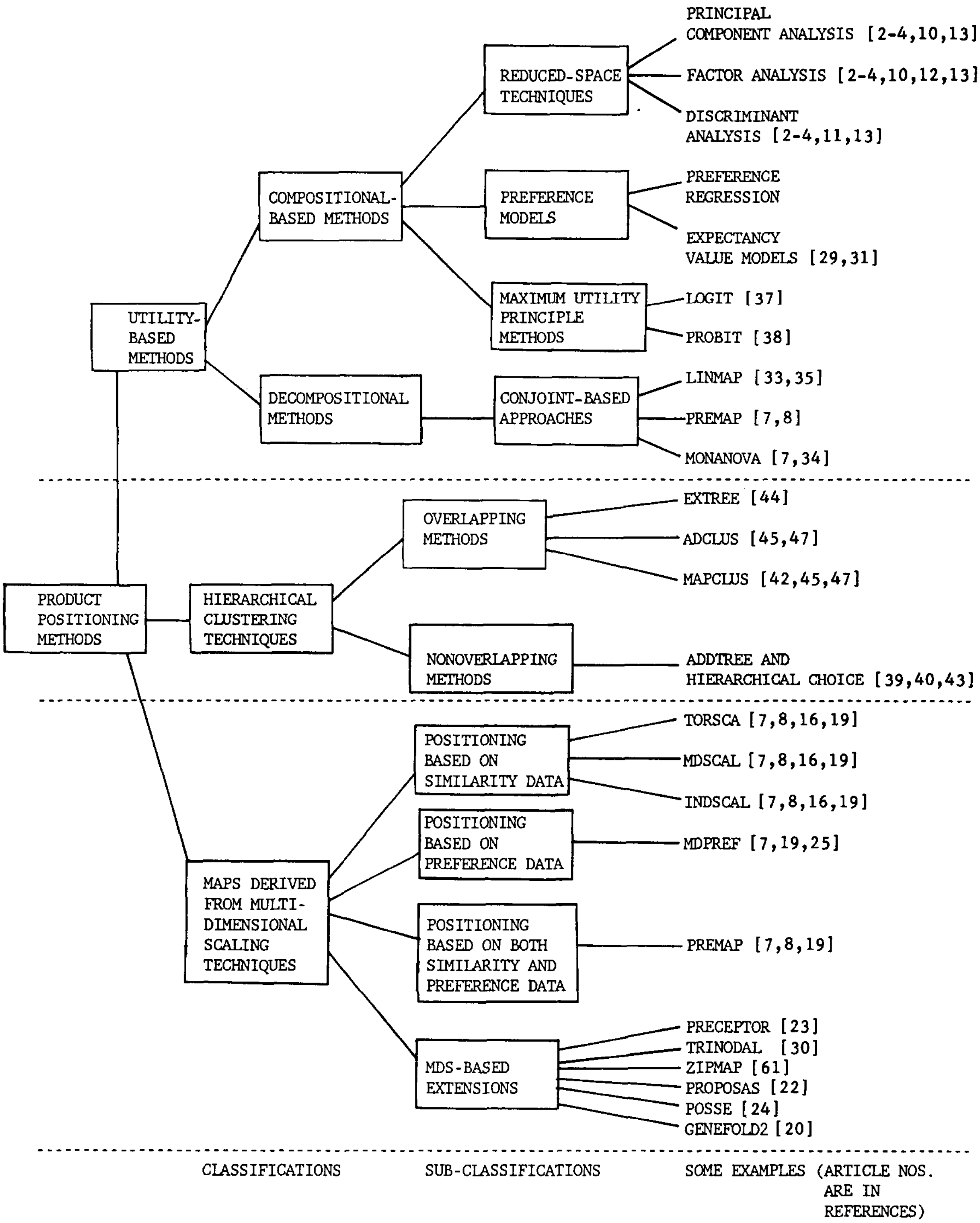

Fig. 1. Product positioning methods.

Data input. The common practice is to use the $n \times p$ (individuals $\times$ variables) matrix of raw data and transform it to a matrix of correlations, using the ordinary product moment transformation. The primary reason for use of the matrix of the correlation coefficient is that the variables under study may have different/arbitrary units and scales. Transformation makes the variables directly comparable. It should be pointed out here that use of the covariances has not been prevalent in PCA applications, although it has desirable properties (see Ref. [13]).

Factor analysis (FA)

FA is the study of interrelationships among the variables in an attempt to find a new set of variables, fewer in number than the original set [13]. Like PCA, FA requires consumers' ratings of all the products as its input. This technique attempts to isolate the underlying factors that best explain the consumers' ratings. FA looks for a correlation between the rating scales and estimates their correlations to the underlying dimensions. These correlations are called factor loadings. Then one can name the underlying dimensions by examining the factor loadings. In brief, the way FA works is that it examines the matrix of correlations among the basic attributes and extracts factors one by one in order of the variance explained.

The primary goal of PCA is to construct linear combinations of the original variables which account for as much of the (original) total variation as possible. In contrast to PCA, in the FA model, interest centers on only that part of the variation that a particular variable shares with the other variables forming a set.

With the aid of these two data-reduction techniques, the analyst may calculate the projection of each observation on each of the factors. In PCA, principal component scores can be calculated directly as a linear combination of the original or the standard variables. In FA, the factor scores cannot be calculated directly, but instead must be estimated. Factor scores give the location of each observation (individual) in the space of the common factors. It is by averaging these factor scores across all individuals that the position of a product on the perceptual map is determined. When we do this for all the factors and all the products, we have a perceptual map.

Multiple discriminant analysis (MDA)

Discriminant analysis is a technique for classifying individuals/objects into mutually inclusive groups on the basis of a set of independent variables. This technique can handle either two groups [13] or multiple groups [11]. In the two-group case, the idea is to find a single linear composite (function) of the independent variables (attributes) that would discriminate between the groups. Essentially, the single linear composite provides a new axis along which the groups were maximally separated.

In MDA the objective is the same in that one finds an axis with the property of maximizing the ratio of between-groups to within-groups variability of projections on this axis. In general, with M-groups and p-independent variables, there are, in total, $\min(p, M-1)$ possible linear composites (axis). In most applications, since the number of predictor variables far exceeds the number of groups, at most, $M-1$ discriminant axes (composites/functions) will be considered.

Determining the number of statistically significant† functions is a particularly important issue. The number of discriminant functions which provide statistically significant among-groups variation essentially defines the dimensionality of the discriminant space. Thus, MDA can be viewed as another data-reduction technique.

In discriminant analysis the mean value of the discriminant function is commonly referred to as the group centroid. Utilizing the statistically significant discriminant function, the group centroids can be plotted in the reduced discriminant function space to exhibit geometrically the separation of groups.

It has become customary to transform the discriminant weights, associated with predictor variables, into what are labeled as *discriminant loadings*. Discriminant loadings are simply correlations, where the correlation is taken across the observed values of a given variable and the computed discriminant scores for the discriminant function. Discriminant loadings play a dominant role in aiding managerial interpretations.

Unlike PCA, which tends to orient a space to dimensions that have high variance both within and between objects, discriminant analysis orients a space to those dimensions that have high variance across objects, but low variance for subjects rating a given object.

†To know how many linear functions to retain, Bartlett's statistic and its χ^2 approximation is commonly employed.

Data transformation in reduced-space mapping techniques is important to remove *bias* in the data. In perceptual mapping applications, the data format generally used is in the form of: respondents × brands × attributes [4]. In perceptual product space analysis, Harshman and Margaret [15] discuss preprocessing data, which refers to the transformation of the data values prior to applying techniques, and before generating the perceptual space. Processing transformations, which have been used in perceptual product space analysis, may be classified as: standardization by individuals [3], standardization across respondents and brands [16] and normalization within individuals [2]. Dillon *et al.* [4] provides an extensive discussion of these data transformations used in perception mapping techniques. The most commonly mentioned reason for the use these three tranformations is to remove "scale-bias" and "yea-saying" effects [2–4, 17]. This type of transformation seems consistent pragmatically as suggested by Harshman and Margaret [15].

Another class of methods, which has dominated current research on product positioning, for determining the perceptual space/stimulus configuration are the multidimensional scaling techniques that are discussed below (see Fig. 1).

Perceptual Mapping/Stimulus Configurations Based on Multidimensional Scaling (MDS) Techniques

MDS is another mathematical tool which is frequently applied to product positioning problems [7, 16, 18–24], and attempts to represent the proximities (similarities) between objects (products) spatially, as in a map. Although MDS is a complex psychometric technique, the principle behind this approach is quite simple. The primary objective of this technique is to map the objects in a space in such a way that their relative positions in the space reflect the degree of perceived proximity (similarity) between the objects.

Unlike the previous data-reduction techniques that utilize brand attribute rating data, MDS is based on consumer judgments or preferences of the brands. In other words, the consumers' similarity/preference data of product pairs is the basic input for an MDS program. MDS methods are typically used to provide discriptions of preferences and/or similarities of a sample of consumers towards a set of items in a product category.

A variety of MDS programs, which can assist in identifying "best" locations for existing of new products in the perceptual space, are now available. In the last decade or so, there has been a rapid increase in the number of computer programs for MDS. The major available computer programs are: MDPREF, INDSCAL, MDSCAL, PARAMAP, PREFMAP, TORSCA, PROFIT AND MINISSA [7, 19, 25].

The most commonly employed algorithms either use similarity data or preference data. The multidimensional algorithms (e.g. TORSCA) that use similarity data construct geometrically spaced models such that the more similar objects (brands) are placed close together.

Preference data algorithms (e.g. PREFMAP) produce joint space maps of both consumer ideal points and brands [8, 26, 27]. These joint space maps are constructed so that consumers' ideal points and objects are placed in a common spatial configuration. For nonmetric maps of this type, the objects and ideal points are positioned so that the object's relative proximity to the individual's ideal point best preserves the rank order preference of that individual.

A detailed discussion of mapping configurations of the various MDS algorithms, based on the data type, will be presented next.

Types of models/algorithms used to portray the stimulus configuration when using different data input in the MDS

MDS methods, considered generally, attempt to represent certain types of data as relations on points on a multidimensional space. The dimensions of the space are assumed to represent attributes or certain properties along with which stimuli (objects) are compared. A variety of models have been proposed for the analysis of similarity, preference or similarity–preference data.

Whether one uses similarity (perception), preference or both, it is important to understand models that portray different stimulus configurations (maps). Various types of similarity (proximity) matrices and algorithms have been developed to portray similarity relations geometrically.

The major model used to represent similarities (or dissimilarities) is the *distance model* [19, 28]. Many distance measures are special cases of the Minkowski metric, defined by

$$w_{ij} = \left(\sum_{k=1}^{K} |X_{ik} - X_{ij}|^r \right)^{1/r}, \tag{1}$$

where δ_{ij} denotes the distance between stimuli i and j and X_{ik} and X_{jk} represent the response on the kth characteristic, $k = 1, 2, \ldots, K$. All the distance models considered in mathematical psychology assert that the similarity between a pair of stimuli is some function of their parital dissimilarities, with respect to each of several perceptual dimensions.

The most commonly used algorithms (e.g. MDSCAL, INDSCAL, TORSCA, PARAMAP) that use similarity data construct geometrically spaced models such that the more similar objects (brands) are placed close together. To mention succinctly, the MDSCAL and INDSCAL generate different results. For instance, MDSCAL represent the products in k-dimensional space such that the Euclidean distance between pairs of products, d_{jm}, best produces the rank order of the judged similarities among the products; the Euclidean distance is measured by (see Refs [7, 29])

$$d_{jm} = \left[\sum_{k=1}^{K} (x_{jk} - x_{mk})^2 \right]^{1/2}, \tag{2}$$

where x_{jk} is the measure of product j's position on the kth dimension as produced by the algorithm.

The goodness-of-fit criterion often called "stress" is calculated using the following stress formulas:

$$\text{Stress(formula 1)} = \left[\frac{\sum_{j=1}^{J} \sum_{m=1}^{J} (d_{jm} - \hat{d}_{jm})^2}{\sum_{j=1}^{J} \sum_{m=1}^{J} d_{jm}^2} \right]^{1/2} \tag{3}$$

and

$$\text{Stress(formula 2)} = \left[\frac{\sum_{j=1}^{J} \sum_{m=1}^{J} (d_{jm} - \hat{d}_{jm})^2}{\sum_{j=1}^{J} \sum_{m=1}^{J} (d_{jm} - \bar{d})^2} \right]^{1/2} \tag{4}$$

where $\bar{d}$ is the arithmetic mean of the estimated distances. The $\hat{d}_{jm}$ are a set of values chosen to be as close to the d_{jm} as possible subject to the constraint that they be monotonically related to the original data. Stress (formula) 1 and 2 differ only in the normalizing constant for the denominator of the fraction under the square root sign. Stress is minimized via the nonlinear programming method of steepest descent.

The mathematics of INDSCAL, however, differs from MDSCAL in that it provides individual i's perception of the jth product along the kth dimension. Then the modified Euclidean distance, d_{jm}, between two products, j and m, for an individual i is given by

$$\delta_{ijm} = M(d_{ijm})$$

and

$$d_{ijm} = \left[\sum_{k=1}^{K} (w_{ik}(X_{jk} - X_{mk})^2 \right]^{1/2}, \tag{5}$$

where

δ_{ijm} = ith individual's similarity ranking for products j and m,
d_{ijm} = weighted distance between j and m,
x_{jk} = coordinates of stimulus j on dimension k,
w_{ik} = weight individual i attaches to dimension k

and

M = any monotonic function.

INDSCAL provides individual level estimates of product perceptions. In this method, instead of pooling similarities data from all respondents, it is assumed that although all share a common perceptual map, individuals may weight the dimensions of the map differently. For details of the algorithms, see, for example, Refs [7, 16].

When preference data is used, the models used to generate the stimulus configurations are the distance and vector models [7, 16, 26–28]. The conceptual notion underlying the distance models in the context of preference is the *ideal point*, a hypothetical stimulus encompassing a specific combination of "scores" on underlying perceptual dimensions. All vector model formulations can be viewed as a special case of the distance model in which the ideal points are assumed to extend to infinity. Carroll and Chang's [25] MDPREF model uses a vector formulation which can handle, metrically, the case of either rank ordered data or paired comparisons.

Alternative methods for dealing with a joint space representation of stimuli and ideal points, using either similarities or similarities and preference data, have been also used extensively in product positioning analysis. There are two approaches to deal with joint space representation: *ideal point formulation* and *ideal point generalization*. In the ideal point formulation, similarity data are collected for all pairs of $(n + 1)$ stimuli (the $n + 1$ stimulus being an *ideal* brand). Here respondents are asked to imagine what their ideal brand would be like when the similarity data are scaled by the usual distance model; the distances of the real stimuli from the *ideal* are assumed to reflect decreasing preference (the maximal preference at the ideal and a utility function that declines monotonically with the distance from the ideal point).

In the *ideal point generalization* approach, which uses both similarity and preference data for the same group of subjects and stimuli, Carroll and Chang [27] proposed models that can be distinguished on the basis of being metric vs nonmetric. The metric version assumes that an individual's utility function is linearly related to the weighted squared distance of each of the real stimuli from his ideal. The metric version of the model proposed is PREFMAP (see Ref. [19]). In Carroll–Chang's nonmetric version of the PREFMAP, the utility is assumed to be only monotonically related to the weighted squared distance from the ideal. The models that develop joint space configurations from preference data alone, of course, obviate the need for collecting the similarity judgments. The trade-off here involves a loss of information regarding the transformation associated with going from a similarities to a preference context, and the assumption that all respondents visualize the stimuli in the same way.

Recently, extensions of joint space mapping routines have been developed. Keon [30], for example, introduced an MDS routine called TRINODAL. The basic advantage of the TRINODAL MDS routine over other mapping techniques is that it simultaneously plots consumer ideal points, ad images and brand images onto a single map. The MDS routine first reorders the ads according to the lowest variance, and then reorders the brands so that the first brands load most heavily on the earlier ads.

The scaling of the ads and brands is done by assuming that the proportion of times and advertisement i is identified as being for brand j can be considered a proximity measure of the relative spatial distance between the ad and the brand in a perceptual configuration. Mathematically it is expressed as

$$\delta_{ij} = d_{ij}^{-2} \Big/ \sum_{k=1}^{m} d_{ik}^{-2}, \tag{6}$$

where

δ_{ij} = the proportion of times advertisement i $(i = 1, 2, \ldots, n)$ is identified as being for brand j $(j = 1, 2, \ldots, J)$

and

d_{ik} = distance of advertisement i to brand k $(k = 1, 2, \ldots, m)$.

Each ad thus establishes a set of the above relationships, one for every brand. Each set of relationships represents a series of concentric spheres in k-dimensional space where a particular brand can be positioned.

Utility-based Compositional Methods Used to Generate Perceptual Maps

The compositional methods, which are based on the utility principle (see Fig. 1), are: the expectancy value models, regression, logit/probit etc. It should be pointed out that the reduced-space techniques presented earlier, which include PCA/FA and the MDA, are also compositional approaches. However, these dimension-reducing methods simply aid practitioners in condensing the data into fewer dimensions using the compositional functional form; where the objective is to reduce data into space of fewer dimensions. In the utility-based compositional approaches, however, the underlying principle is to choose an object with the highest utility rather than reducing the data into fewer dimensions. Compositional techniques, in selecting the best positioning of a product, the best physical features to achieve that positioning and which product or set of products is best to cover the market, are the preference models (Fig. 1). By using the preference models, an evaluative tool, a manager can predict how consumers will prefer each positioning and thus make a selection amongst a number of alternative product positions.

Preference models

Perhaps the easiest techniques to measure consumer preferences are the *expectancy value models* (*EVM*). EVM asks consumers to directly specify product perceptions and the importance weights of the attributes embodied in the product. The preference of individual i for product j is given by

$$p_{ij} = \sum_{k}^{p} W_{ik} X_{ijk}, \tag{7}$$

where

w_{ik} = the importance individual i places on attribute $k (k - 1, 2, \dots, p)$

and

X_{ijk} = individual i's perception of product j relative to attribute k.

The model, equation (7), has been used extensively to obtain consumer preferences [29, 31].

The expectancy value model uses "self-stated" importance weights given by consumers. Though these weights are indicative of what consumers believe to be important, one cannot assert that they capture the range and scale used to mathematically represent the attributes as perceptual dimensions. One way to capture these effects is to use a statistical model that *estimates* the importance weights and *best* fits the observed preferences. Preference regression is one of these techniques.

Preference regression analysis (*PRA*)

PRA is a statistical technique that produces *average* importances for the dimensions of the perceptual map [29]. These average importances and the position of the products on the map help identify positioning opportunities. In PRA, the model used is similar to the self-rated importance model used to represent preference. However, there are subtle differences. These are: first, in preference regression analysis, perceptual dimensions are used rather than the basic values of the attribute ratings to measure consumer perception; second, importance weights are estimated rather than asking the consumers to state them directly.

These methods are powerful techniques for analyzing consumer preferences. The EVM offer an inexpensive way to get an idea of the linear effects of product attributes in formulating preferences. PRA provides a method of estimating the importance of psychological dimensions used to define perceptual maps. Other compositional methods, such as logit/probit, which are probabilistic in structure, are discussed under the rubric of estimation methods for conjoint analysis, which relates paired-comparison data to a choice probability model.

Utility-based Decompositional Methods for Product Positioning Analysis

The underlying principle here is that decomposition models express the overall value of each object (product) as a function of the real value associated with its components. This class includes all variations of the expected utility theory, as well as the various adding and averaging methods. The most widely used method is conjoint analysis [32].

Conjoint analysis (CA)

CA permits determination of consumer part-worth utilities for different attributes at their levels. An understanding of such part-worths would enable practitioners to decide: (i) what should be the basic features of the product they offer and (ii) at what level should these features be offered?

In CA, the respondent is asked to react to a product alternative in terms of an overall evaluation. The overall evaluation is generally a preference rating (e.g. the respondent may be asked to rank the alternative products by overall preference). There are two basic data collection procedures used in CA: (i) the trade-off procedure and (ii) the full-profile approach. The trade-off procedure is also commonly referred to as the "two-factor at a time" procedure [33]. For a detailed discussion of the advantages/disadvantages of the data collection approach, readers may refer to the Green and Srinivasan [32] article.

CA is thus a mathematical technique used to outline the ranking information that is useful to managers, in order to understand the product features and thereby make positioning/designing decisions. The conjoint estimation equation is given by

$$F'_{ij} = \sum_{k=1}^{K} \sum_{l=1}^{L} U_{ikl}\delta_{jkl} + \text{error}, \tag{8}$$

where δ_{jkl} is the indicator variable which tells us whether product j has features k, at level l; and is given by

$$\sigma_{ikl} = \begin{cases} 1, & \text{if product } j \text{ has feature } k \text{ at level } l \\ 0, & \text{otherwise.} \end{cases}$$

U_{ikl}'s (called part-worths) represent the utility level that is appropriate to show individual i's valuation of having the kth attribute at the lth level.

In CA, there are various approaches to computing the utility scale of each attribute. These utilities indicate how influential each attribute is in the consumer overall evaluation. Techniques for estimating utility scale values differs depending upon whether the preferences of the input data are assumed to be ordinal (rank order only) or interval (ratings). Among the class of techniques that treat the preference data as, at most, ordinally sealed are: MONANOVA [34], PREFMAP [8], LINMAP [35] and Johnson's nonmetric trade-off procedure [36]. The most widely used method when the preference data are assumed to be intervally scaled is ordinary least-square regression [14]; a succinct discussion of these methods is presented below.

Estimation methods for part-worth utilities

Kruskal's [34] MONANOVA (MONotonic ANalysis Of VAriance) algorithm performs an additive main-effect ANOVA on monotonically transformed values of the original data. As the name suggest, in the MONANOVA, the response variable needs only to be rank ordered. MONANOVA attempts to find a monotone transformation of the input data whose values are represented as an additive combination of main effects. The monotone transformation is obtained on an iterative basis that seeks to minimize what is known as Kruskal's stress [34].

LINMAP, developed by Srinivasan and Shocker [35], is another approach to estimate part-worths in CA. The major advantage of this approach is that it uses a more robust error structure and that constraints can be added to ensure certain properties of the utility function. In the LINMAP procedure, the consumer is asked to rank order the products, and from this a rank order among all pairs of products is derived. To illustrate this, using equation (8), one can estimate part-worths U_{ikl}, as follows:

$$\min \sum_{j=1}^{m} \left| F'_{ij} - F_{ij} \right|$$

subject to

$$F_{ij} = \sum_{k=1}^{K} \sum_{l=1}^{L} U_{ikl}\delta_{jkl} \qquad (\forall j);$$

and

$$U_{i,k,l+1} \geqslant U_{i,k,l} \qquad (\forall k, \text{ for } l = 1, 2, \ldots, L-1)$$

and

$$U_{ikl} \geqslant 0 \qquad (\forall k, l).$$

Here it should be noted that the first constraint defines the linear form and the second constraint ensures a monotonic utility function. The final constraint simply defines the scale for the utility function and has no behavioral meaning. Since LINMAP is very flexible, it has been gaining wide popularity. For a detailed discussion of this approach, refer to Refs [29, 35].

Other methods for estimating part-worth utility values are: ordinary least squares (OLS), and stochastic modeling approaches such as: logit [37] and probit [38]. The OLS procedures attempt to minimize the sum of squared deviations between the observed and predicted preference values. Stochastic models such as the logit/probit maximize the likelihood function to estimate part-worths. The basic rational behind the logit/probit analysis is a mathematical function that links preference values to choice probabilities, based on the theory of consumer behavior developed by McFadden [37].

Another approach that has been frequently used in product positioning is cluster analysis. Like MDS, one can analyze proximity data using hierarchical cluster analysis. Both the methods are built on distance models. Despite the similarity between these two approaches, they fundamentally differ. The relationship between the proximity data and the distances in cluster analysis often cannot be expressed by a linear or even a monotone function as in MDS. Further, cluster distances are not spatial distances as in MDS.

One may further divide clustering approach into nonoverlapping or overlapping clustering; each of these approaches are next discussed.

Nonoverlapping (Hierarchical) Clustering Analysis

Evidence indicates that consumers simplify a problem according to a hierarchical structure/scheme; whereby the total set of items is sequentially partitioned until a subset of items (which are considered as substitute alternatives) exists [39–41]. Hierarchical clustering approaches are generally based on the computation of distances between pairs/clusters of objects or brands; then at each clustering stage, objects that are closer together are grouped first. Computational algorithms offer choices among alternative methods (e.g. minimum average or maximum linkage [42], which can be used to identify points between which the distance is then computed; for example, by obtaining the sum of squared distances (d^2) between pairs of points (objects or clusters) or some characteristics (e.g. product attributes). This distance may be represented as [40]:

$$d_{\mathrm{t}}^2(\text{total}) = d_{\mathrm{a}}^2(\text{accountable}) + d_{\mathrm{c}}^2(\text{chance}).$$

Thus, if the "total" similarity measure is used as input into a clustering algorithm the resulting structure is subject to the optimization of the chance variation. This approach is based on the concept that given a relevant set of variables, objects should be grouped in an iterative† manner so as to maximize the prediction of the accountable variance between objects on some relevant criterion available.

In the psychology literature, Sattath and Tversky [43], for example, developed a computer algorithm, labeled ADDTREE, for the construction of additive similarity trees. Its input is a symmetric matrix of similarities (or dissimilarities), and its output is an additive tree. In the present algorithm the construction of the proceeds in stages by clustering objects so as to minimize the

†The iterative grouping procedure is conditional on previous groupings, just as predictor variable selection in stepwise regression is dependent upon variables already selected.

number sets in satisfying the *additive inequality* which is given by

$$d(x, y) + d(u, v) \leqslant \max[d(x, u) + d(y, v) + d(x, v) + d(y, u)], \tag{9}$$

where x, y, u, $v \in S$, and S is any set of objects.

A dissimilarity measure d on S is a nonnegative function $S \times S$ such that

$$d(x, y) = \begin{cases} d(y, x) & \\ 0, & \text{if } fx = y. \end{cases}$$

Sattath and Tversky [43] point out that the additive inequality is both necessary and sufficient condition for the representation of a dissimilarity measure by an additive tree. An ADDTREE is less restrictive than the above discussed hierarchical clustering scheme; and, as noted by Tversky and Sattath [43], it has some empirical and theoretical advantages over the MDS.

Users of the hierarchical clustering algorithm, such as discussed above, tend to assume that clusters are disjoint; i.e. clusters are nonoverlapping. Users of these approaches, however, tend to discard much of the detail (e.g. levels of nesting for specific clusters) found in the dendrogram (tree structure). The most commonly employed alternative is to obtain a partition of the set of entities being clustered. That is, the objects are segregated into mutually exclusive (and exhaustive) clusters. However, in reality it is easy to find examples in which clusters overlap.

There are techniques such as EXTREE [44] which produce hierarchical structures (trees) that can accommodate overlapping clusters. Similarly, there is a technique called ADCLUS/MAPCLUS [45] which, when used in association with MDS, can produce overlapping clusters on a map representation.

The Overlapping Clustering Concept

The development of overlapping clustering [42, 46] permits researchers and analysts to cluster brands into nonmutually exclusive and exhaustive categories. Conceptually, it is desirable to extend the concepts of product positioning to encompass the cases of overlapping clusters. Methodologically, this extension would require a clustering method to represent the overlapping structure in the data in a parsimonious manner, focusing only on necessary cluster overlaps that can be substantively justified and often spatially presented.

Though the methods of overlapping clustering have been available for decades [e.g. 42], recent methods of product positioning, proposed by Phipps *et al.* [45] and Arabie and Carroll [47], provide new insights into positioning concepts. These approaches are ADCLUS (ADditive CLUStering), and MAPCLUS (MAthematical Programming CLUStering).

The ADCLUS model and MAPCLUS algorithm

Assume that M objects (brands) are to be clustered, with input data of $n(n-1)/2$ entries contituting a two-way symmetric proximity matrix having no missing data.† Let $S = \|S_{ij}\|$ be transformed proximities, with which the fitted S matrix is being compared. The underlying structure of ADCLUS can be expressed as

$$\hat{S}_{ij} = \sum_{k=1}^{m} W_k P_{ik} P_{jk}, \tag{10}$$

where $\hat{S}_{ij}$ is the theoretically reconstructed similarity between object i and j, W_k is the nonnegative weight representing the saliency of the property corresponding to subset k and

$$P_{ik} = \begin{cases} 1, & \text{if object } i \text{ has the property } k \\ 0, & \text{otherwise.} \end{cases}$$

Thus, there are m subsets or clusters of the n objects, and these clusters (which are to be *recovered* or fitted by the clustering algorithm) are allowed to overlap (although there is no explicit requirement that they do so). It should be mentioned that the weights and clusters are both fitted in applying the ADCLUS model. The underlying rationale for the ADCLUS is that the predicted

†The raw data may be in the form of either similarities or dissimilarities; it can then be first transformed linearly to be similarities as the interval [0, 1].

similarity $\hat{S}_{ij}$ of any pair of objects is the sum of the weights of those clusters containing both the objects and i and j.

Arabie and Carroll [47] devised a MAPCLUS algorithm to fit the ADCLUS model. The most obvious difference between the ADCLUS and the MAPCLUS programs is that, for the latter, the (fairly small) number of subsets m, is specified by the user. In practice, MAPCLUS has been able to obtain solutions acceptable in terms of both interpretability and goodness-of-fit, using considerably fewer clusters for various data sets than was possible with the ADCLUS program.

In brief, overlapping clustering can be applied to market analysis (e.g. segmentation) to cluster products (or individuals). The acceptance of this concept, coupled with the likelihood of the presence of both overlapping positioning and segments, suggests the need to explicitly examine the relation between the two sets of results.

3. DECISION ISSUES RELATED TO PRODUCT POSITIONG MODELS

Selection of the Number of Dimensions Based on A Priori Knowledge of the Dimensions Upon Which a Decision is Made

The set of product attributes provides the basic dimensionality for the mapping techniques and also clearly plays a crucial role of theoretical importance. An important characteristic of the mapping models such as: EVM, conjoint, logit, probit and PRA, is the underlying assumption that dimensional inputs are *a priori* known and measured by the researcher. Since, in these procedures, the data is not condensed to obtain the underlying dimensions, the researcher's knowledge regarding the problem is crucial. Several issues surround attribute generation and inclusion of attributes in the model. As Wilkie and Pessemier [31] pointed out, attribute generation operates at two levels: the *initial specification* of attributes in data gathering; and the *inclusion* of *attributes* in the model, which can reflect either a direct use of the raw data, or be the result of reworking the raw data in some manner. A fundamental criterion required for specification of attributes is that they be exhaustive, semantically meaningful, and subject to unidimensional interpretation [31]. Issues concerning the inclusion of the number of attributes, as noted by Wilkie and Pessemier, are: attribute independence, salience vs the importance of the attributes and the minimum number of attributes descriptive of attitudinal structures.

Another important decision issue concerns the number of attributes influencing the choice decision, and also the homogeneity of the population. Within the sample (population), individuals differ in their background variables, such as: education, familiarity with the products, past experience etc. Therefore, different individuals in the population may use a different number of attributes when they make their decisions. Thus, by segmenting the population on the basis of, for example, background variables, a different set of variables may be included in the model to generate perceptual space.

Determination of the Dimensions Which are Not Known A Priori

Models, such as the PCA/FA and discriminant analysis, enable researchers to generate dimensions directly using raw data. Here it is important to realize that the derived perceptual space (i.e. the "factors/dimensions") generated by the use of a data-reduction procedure such as the PCA/FA, is actually a functin of the raw data collected. In the data collection procedure, if one attribute is asked for in several different forms, that attribute is more likely to constitute a major dimension of the derived perceptual space. In other words, the number of dimensions generated may be more a function of the attributes asked for rather than of the product features that consumers perceive as important [4, 48]. As extensively discussed by Schocker and Srinivasan [48], if the analysts wish to obtain the product space in actionable/interpretable terms, then the attributes selected must not only be meaningful to consumers, but also be easily controlled by marketers. Furthermore, care must be taken in defining the underlying dimensions on which products are scaled (prior to using PCA/FA), and in relating these dimensions to modeling consumer choice (or preference). The argument whether preference/choice is truly a function of the original attributes or of "latent" attributes that are uncovered through reduced-space techniques is debatable. If it is a function of the original attributes, care must be exercised in scaling them so that they will load on the reduced-space dimension in the same way that they influence preference/choice. Pragmatically, however, this is not always feasible.

Compared with dimension-reducing methods, MDS requires no prespecification of attributes. Consequently, MDS may serve as a tool to check on the set of attributes that can distinguish the products in the space and can be associated with customer preference/choice. Each methodological approach, no doubt, has certain advantages and disadvantages. Some researchers [e.g. 49] suggest working directly with respondents' attribute perceptions, while other researchers [23, 50] propose to use PCA/FA, MDA or MDS to determine a configuration in reduced space.

Issues Related to Levels of Aggregation

Research decisions involved with product positioning studies deal with three types of analyses: the individual level, the segmented level and aggregate level modeling analysis. In conducting a positioning study, the researcher can choose the desired level of modeling based on his assumptions and the type of analysis required. The purpose here is to briefly explicate the strengths/weaknesses of these approaches.

Individual level models

The early empirical applications of the individual models (e.g. CA) indicated that the predictive ability of the individual level models was very good [7, 51, 52]. When one employs these models, a separate function needs to be estimated for each individual. Therefore, individual level analysis requires that enough information be collected from each individual to estimate separate functions. Moore [53] points out that, though individual models may demonstrate good predictive power, the output of the estimation procedures (e.g. a separate set of importance or utility weights for each individual) makes managerial analysis and understanding more difficult when the number of respondents is large. Moore [53] further notes that several good reasons, which includes data collection, knowledge of the attribute levels, time, cost etc., can be given for not analyzing the data at the individual level; however, this needs to be balanced against the very good power of the individual models.

Aggregate models

At the other extreme of the aggregation continuum is the case in which the information is pooled across all respondents and one function is estimated. The results appear to be easy to explain (e.g. on average, convenience is more important than trip cost in determining the transportation mode). However, a potential problem associated with pooled analysis is termed "the majority fallacy" by Kuehn and Day [54]. The majority fallacy is caused by heterogeneity of preferences; for example, if half of the people like large cars and the other half like small cars best, the "average" person may like medium-sized cars best, even though no real person wants one. This problem has been demonstrated in CA settings by Huber and Moore [55].

There are other subtle issues involved in using the aggregate approach. Dillon *et al.* [4] discuss the effects of aggregation on perceptual space. They point out that if PCA/FA is used in uncovering latent dimensions (i.e. factors) that define reduced space, interattribute correlations are typically computed by stacking the brand-by-attribute subarrays for all the respondents into one long matrix. They contend that this approach, in which correlations are calculated across an entire sample (i.e. respondents and brands), can produce misleading results. The primary reasons, as they pointed out, are as follows.

First, the correlation between the attributes is a function of two different sources: one source, which contributes the correlation, relates the consistency/inconsistency of responses to the set of attribute characteristics. Another source relates to how homogeneous or diverse the respondents are in their attribute rating judgments.

The second reason is related to pooling attribute ratings. An approach in which interattribute correlations are computed across an entire sample produces essentially an *average* correlation coefficient. Such average correlations can be acutely affected by mean differences introduced by brands, by *outliers*, or untypical responses. There is also empirical evidence to suggest that pooling attribute ratings across respondents and brands is not to be recommended [56]. Thus, there is convincing evidence to question the common practice of reporting aggregate product space analysis that do not retain differences due to individuals, or brands.

Segmented models

One would then like an approach that combines the most desirable properties of the two extreme levels (individual and aggregate) and avoids the problems of each. Wind [5] cogently notes that there is a strong relationship between segmentation and positioning and the two decisions must be made jointly.

Several interesting approaches towards the creation of segments appear promising. Pessemier [57], for example, suggested using dollar metric sealed preferences as the basis for forming "product preference market segments". Each segment would be composed of individuals who had common relative preferences for all existing products. Urban [23], on the other hand, suggested defining segments on the basis of "similar" distances between the brands and the individual ideal points. Such a procedure might result in fewer segments than would be obtained from the Pessemier approach. If one attempts to segment on the basis of ideal point locations alone, or for that matter, attribute salience alone, these procedures offer less desirable criteria than those mentioned above.

In a review of market segmentation, Wind [58] refers to two traditional methods of segmentation, *a priori and clustering*, and two newer methods, flexible and componential. The clustering and componential methods, which are also mentioned by Green and Srinivasan [32], are frequently used techniques in product positioning. Cluster (segments) are formed by grouping the respondents into segments that are homogeneous with respect to the benefits they seek from the product/service class. Green and Srinivasan [32] suggest they have achieved better discrimination by clustering on the vector of part-worth utilities, rather than on the most preferred level of each attribute.

The componential segmentation approach focuses on the effect of the interaction between the product profile and the persons' profile. Through this mechanism one is able to predict how a person with a certain set of background characteristics will react to a particular product. A potential advantage of the clustering approach over componential segmentation is that people in different segments need not differ in terms of background variables. Componential segmentation will uncover the groups only if they differ in terms of background variables. It must be also mentioned that to assess the key discriminating characteristics, reduced-space techniques such as the discrimination analysis approach has been employed.

Interpretation of the Diagnostic Information Generated from the Output

Perceptual maps that come in different forms generate different types of diagnostic information. This information may be related to finding the underlying "latent" attributes (factors), estimated values of the part-worth utilities, estimation of the ideal point and distance between brands etc. For the consumer analyst to correctly understand the output is imperative. The consumer analyst may fact such situations as retaining the number of dimensions, naming the dimensions and/or interpretation of the solution obtained from the perceptual mapping procedures; each of these issued will be discussed next.

Decisions to retain the number of dimensions

When one employs reduced-space techniques to produce a product space, such as FA, there are two classical rules of thumb for determining the number of dimensions (factors): The "scree test" and the "eigenvalue" rule. In the former case, one plots the incremental variance explained by each factor vs the order in which the factors were extracted. Factors are retained up to the point where the incremental contribution levels off. The latter rule is akin to the scree test, except the cut-off is known when the fraction of variance explained by the next successive factor drops below $1/p$; where p is the number of attributes. In practice, either of these rules are used to retain the appropriate number of dimensions. Recent advances, however, have shown some sophisticated procedures for determining the number of factors [59]. While working with FA, once can name the dimensions by examining the factor loadings.

Like FA, MDS gives analysts the positions of the products on the underlying dimensions. A decision regarding how many dimensions to retain may be made simply by choosing the dimensionality with the smallest stress value. Here one can plot the stress value (S) against the number of dimensions (P). The stress value decreases as the number of dimensions increase. When S is plotted against P, the resulting marginal changes in stress can be examined in the same way as in FA, except that we are examining the stress value rather than an increase in incremental explained variance.

Interpretation of the stimuli in the derived space is very important as it reveals the underlying position with competing products; this is discussed under the rubric of solution interpretation.

Solution interpretation

To interpret the stimulus configuration, one can either simply use the configuration itself (the position of the stimuli in space) to determine the interpretation (the subjective approach), or undertake a more objective approach. The subjective approach to interpreting the configuration rests on the position of the stimulus objects in the stimulus space. The procedure here is to look at the properties (e.g. diet vs nondiet, light vs heavy etc.) of the stimuli occupying extreme positions in the configuration (map). However, it should be noted that the problems encountered in the interpretation of similarity configurations concern the appropriateness of the dimensional (sweer vs unsweet) interpretation itself. In the context of joint space configurations of stimuli and ideal points (or vectors), the problem concerns the interpretation of configurations as obtained from the preference data alone, and their correspondence to the interstimulus distance obtained from the similarities data.

Several objective methods have been used to assist researchers in this task. They include "property fitting" procedures and experimental design methods. For a discussion of these approaches, the readers are referred to Ref. [19].

4. FACTORS INFLUENCING THE TYPE OF MODEL TO USE

Knowledge of the Problem Area

The knowledge of the researcher plays an important role in selecting the appropriate technique. For example, if a consumer analyst has knowledge of the attributes (e.g. price) and their levels (e.g. \$8, \$10, \$11) and is further interested in new product positioning, then he may prefer to choose the conjoint approach. If the analyst believes that the population is not homogeneous with respect to the choice process, then he may segment the population and thereby wish to use different mapping procedures. Based on his knowledge of the problem area, he may choose a data-reduction technique to reduce the problem to a few dimensions and then choose a perceptual space technique to obtain the product space analysis. Hauser and Koppleman [3] point out that an analyst has neither the time nor the money to simultaneously apply all the techniques. He/she has to select one method and use it to address the particular marketing problem at hand. Furthermore, they suggest that the use of the technique should parallel as closely as possible the recommended and common usage. In essence, if the researcher/analyst is knowledgeable about the problem area, the decision as to which technique to choose primarily depends upon whether the added expense in gathering more information is worth the added insight obtained from the analysis.

Type of Data Available

Availability of the data is another factor that influences the choice of technique. The data types used to generate perceptual maps broadly are divided into two approaches. One approach in the construction of the maps utilizes the respondents' evaluation of the products on a set of predetermined attributes in a series of scales. These scales represent the way products are evaluated in the market place. Another approach to the generation of the stimulus configuration depends on the nonattribute-based data; namely, similarity (or proximity) measures simply reflect the perceived similarity of two objects in the eyes of the respondents. The scale ranges typically from "extremely similar" to "extremely dissimilar". The number of pairs to be judged for the degree of similarity is equal to $n(n-1)/2$, where n is the number of objects. Preference data also contains similarity information. Here respondents provide direct judgments of similarity between brands measured by the rank order information. In general, a respondent is expeceted to rank close together those objects perceived to be similar. If the preference information were obtained not as rank order but rather on an interval scale, this information could yield a similarity measure between objects.

If one has brand attribute rating available, reduced-space techniques such as PCA/FA or MDA may be used to construct perceptual maps. If the desire to determine which attributes/features discriminate between brands, MDA may be employed [33]. Alternatively, using brand attribute data, FA may also be utilized, which provides *factor scores* across all individuals. In averaging these factor scores, a perceptual map is determined [10, 29]. In this attribute-based approach, if information regarding the levels of each attribute is available, CA may be used.

If the data is collected in the form of similarity measures, the MDS technique may be utilized to produce perceptual maps. The important point here is that, based on different data inputs, different perceptual maps can be produced, which thereby yield different managerial diagnostics. In brief, data input is an important issue in selecting the appropriate mapping method. Furthermore, the type of data available has an effect on the perceptual space in terms of data aggregation, data transformation [4], the method of obtaining or estimating parameters (e.g. ideal points) of interest, and managerial usefulness [3].

Purpose of the Study

The primary objective of the research is to develop a taxonomy which will indicate what type of technique should be used. The research problems may be to position/reposition and/or design/redesign existing or new products. Based on the research objective, the problem is formulated, the existing methodology may be employed, a new approach can be developed or the present technique is modified.

Let us suppose that the research problem is to determine positioning based on the similarity of brand j to other existing brands. In this case, a direct measure of the similarity data is submitted to the MDS programs to produce a perceptual map. For instance, given proximity data defining the rank order of similarities or dissimilarities on a set of stimulus pairs, the TORSCA program may be used to find a stimulus configuration. TORSCA finds points (objects) in a specified number of dimensions such that the rank order or interpoint distances best reproduces the original rank order of input data.

The researcher can choose the desired level of analysis based on his assumption as to the homogeneity of the market's perceptual configuration. If the researcher assumes that respondents have common perceptual dimensions, however, they differ with respect to the weights they attach to the various dimensions. In this approach, one can use the INDSCAL model algorithm [60].

In the above TORSCA and INDSCAL approaches, the perceptual configuration is based on the consumers' similarities (perceptions) of products; the consumers preferences were ignored. If the researcher wishes, he may use both perceptions and preferences for positioning analysis. The data for this approach may be a simple rank order of the objects according to consumer preferences. A joint space program such as PREFMAP can be utilized, resulting in a joint space of the original similarity space and "ideal" points. The closer a brand is to the "ideal", the more that brand is preferred.

MDS methods, discussed earlier, typically provide descriptions of preferences and/or perceptions of a sample of consumers toward a set of objects in a product category. These methods, no doubt, assist in identifying the "best" locations for existing or new products in the derived perceptual space. However, if the researcher is interested in such issues as optimal product positioning and/or optimal product design, an understanding of extensions of the existing or new techniques related to optimal product partitioning and/or optimal design is imperative. Our attention here is focused on presenting succinctly the mathematical tools that aid in optimal product positioning and design.

Optimal product positioning means to determine the levels of the attributes of a new product to be introduced into a market; such as to meet certain objectives (e.g. profits, revenues etc.) specified by a firm. The optimal product design refers to procedures for not only determining (estimating) the value of the objective function for each point location of interest, but also is a way to search the space systematically to find the specific location that meets the derived objectives.

Recently, several approaches towards understanding positioning/repositioning and design/redesign precedures have evolved; they include: MDS methods, as implemented in the PERCEPTOR model [23]; the TRINODAL model [30]; product positioning using an adaptation of Lancaster's [6] theory of consumer demand; and CA methods [32].

In the PERCEPTOR model [23], physical and psychological product attributes are linked to the trial and repeat probabilities through MDS procedures; and the product design process determines the new brand's position in the perceptual space. The distance from the new brand to the ideal brand specifies its probability of purchase. The PERCEPTOR model provides managers with a better understanding of the perception, preference and purchase structure of their markets, and aids in channeling their creative efforts towards developing successful new product designs.

The TRINODAL procedure analysis data on consumer preferences to simultaneously plot brand images, images and consumer preferences on a single map. This approach, however, lacks directions for product positioning or design.

The CA approach [32] and its extensions directly analyze preference data in terms of product attributes, and estimates (using procedures such as LINMAP, MONANOVA etc.) part-worth functions. These functions are then utilized to determine optimal product (re)designs. Various other procedures have been utilized for product design optimization, and these include: modified gradient and grid search procedures for identifying new product ideas [49]; ZIPMAP [61]; PROPOSAS [22]; and a mixed integer nonlinear programming approach for optimal product positioning proposed by Gavish *et al.* [62] and POSSE [24].

ZIPMAP [61] uses a zero-one integer and nonlinear programming for finding the best perceptual attributes of a new product; PROPOSAS, on the other hand, finds the best profile of perceptual attributes of a new product using the ideal point formulation. POSSE [24] presents a comprehensive CA methodology for optimal product design. More recently, May *et al.* [63] have proposed an interesting nonlinear programming algorithm call QRMNEW, which can incorporate technological range constraints on the product attributes, as well as more general linear constraints. However, this method does not guarantee a global optimum.

Except for the POSSE approach, these optimal product design methods deal with perceptual attributes. In contrast to these approaches, DeSarbo and Rao [20] present a new approach which employs the method of combinatorial optimization to obtain a solution to optimal product positioning and design. This methodology is called GENFOLD 2 (GENeral unFOLDing analysis of preference data—version 2).

The attractiveness of this methodology lies in its structure, which enables the researcher to "manipulate" the derived space in answering various questions about product positioning and/or positioning, optimal product design and targeting products to the designated market segments. The underlying mathematical structure of the GENFOLD 2 model is

$$\Delta_{ij} = \mathrm{a}_i f_{ij} + \mathrm{b}_i + e_{ij} \tag{11}$$

and

$$f_{ij} = \sum_{i=1}^{T} (x_{jt} - y_{it})^2, \tag{12}$$

where:

Δ_{ij} = the "dispreference value" (inversely related to preference values) the ith subject has for the jth stimulus ($i = 1,2,\ldots,I$ subjects; $y = 1,2, \ldots,J$ stimuli); $\mathbf{\Delta} = ((\Delta_{ij}))$;
y_{ij} = the tth coordinate of subject i's ideal point ($t = 1,2,\ldots,T$ dimensions); $\mathbf{Y} = ((y_{ij}))$;
x_{jt} = the tth coordinate of stimulus j; $\mathbf{X} = ((x_{jt}))$;
a_i = subject i's multiplication constant;
b_i = subject i's additive constant;
f_{ij} = squared distance between subject i and stimulus j;

and

e_{ij} = error.

The structure of the above model postulates equal weights for the dimensions; further, individual differences in preferences are modeled by allowing the parameters a_i and b_i to vary among the subjects. Technical details and an estimation procedure may be found in Ref. [20].

The technical details available from the GENFOLD 2 model provide interesting insights into decision making for optimal product(s), design, product positioning/repositioning and market segmentation. Compared to other MDS algorithms, GENFOLD 2 provides an intriguing product positioning map of both the ideal points and the stimuli in a common dimensionality; however, it has some limitations (as pointed out by the authors). One potential problem as recognized by the authors is that GENFOLD 2 does not offer a specific theory to indicate when physical attributes will predict perceptual locations of products, or when personal characteristics will adequately predict individual ideal point locations. Another limitation of GENFOLD 2 concerns the static market assumption. It is assumed that competition is not active during the targeting phase of the optimization. However, the authors point out that these concerns are under investigation.

In addition to the above methods for designing/redesigning products, hierarchical structuring (or tree analysis) reveals some intriguing information regarding product design. Hierarchical structuring analysis posits information on what characteristics/features are guiding the consumer to perceive one brand differently from other brands. When the goal of the manager is product (re)design, the product perceptual mapping techniques may not yield adequate information which would be imperative to the product design. There are, however, instances where employing both maps and hierarchical clustering can enable us to better understand the underlying market structure. Urban and Hauser [29] illustrate an application of PRODEGY to a specific market (coffee market) in which brands were represented by maps, but the substitutabilities of various types of coffee markets were portrayed by a hierarchical analysis. This discussion therefore posits that, depending upon the managerial goal, which may be product (re)positioning, product (re)design, or both, a map, a hierarchical tree, or both, may provide enough information to meet the goal.

5. RELIABILITY AND VALIDITY TESTING

Green *et al.* [24] note that to date, little is known about the reliability and validity of the MDS-based procedures. As far as the reliability of the conjoint methodology is concerned, evidence indicates that as long as the stimulus descriptions are relatively simple and the number of separate evaluations is not excessive, reliability of the conjoint tasks appears acceptable. Furthermore, the internal validity (i.e. estimating judgments for the holdout stimuli during the survey) of the conjoint methods, for example, compares favorably with survey-based techniques in general. As far as the predictive (external) validity is concerned, the predictive ability of the conjoint procedures to predict actual market choice behavior or, at least, judgments made after the survey or experiment, has results that seem to be encouraging [64].

Data used for the reliability test also provides a method of cross-validation. The parameters of the population from the first set of data can be used to predict choices/preferences for the second set. The procedure can then be reversed by predicting from set two to set one, thus completing a double cross-validation. While the internal validity tests the goodness-of-fit of the model, cross-validation also takes into account the predictive ability of the model.

It has become customary to use the double cross-validation procedure on attribute-based methods; such as the PCA/FA, MDA, logit, probit etc. Since the data is collected by survey-based designs (questionnaire designs), it becomes convenient to perform the cross-validation procedure. One can also use statistical tests to test the significance of the predictive ability of the models (e.g. proportion of correct predictions). Use of the statistical test suggests how well the models perform, and which model predicts better than the other.

Published empirical work also provides some additional considerations for the testing reliability/validity. Hauser and Shugan [65], for example, have proposed measures of intransitivity for interval and ratio-scaled preference data. These authors suggest that consistency can be tested at the aggregate level in terms of market shares. Similar approaches have been proposed by other researchers [66]. Scott and Wright [67] recommend some additional consistency checks to test whether the estimated parameters make sense. First, the signs of the estimated parameters should agree with the *a priori* notion, based on the prior theory or reasoning. Second, the parameters derived for different subpopulations should differ in the direction that would be expected from the prior theory. In addition, the face validity (or content validity) of the results can be checked by comparing a respondent's subjective values (self-reports) with estimated parameters. In brief, a reliability and validity check is an important part of gaining a degree of confidence and accuracy about the techniques used. Clearly, product perceptual mapping procedures are only as good as the quality of the basic data that underlies them. Much more research is needed for measurements of reliability and validity.

6. CONCLUSIONS

The more traditional utility value models of product positioning assumed that the researcher knew the full set of relevant variables the customers used to evaluate the alternative brands or products. The models would then indicate the relative salience each independent variable had on preferences and/or purchase decisions. CA is often used when the researcher assumes he knows the relevant independent variables and is interested in how customers evaluate and make relative trade-offs on various *levels* of each independent variables.

The mapping approach with its many MDS algorithms is particularly useful when the researcher is unsure of what are the key dimensions used by customers in evaluating alternative brand choices. As the rate of technological innovation, image creation and multiple-usage products increases, it becomes increasingly difficult to specify exactly what is the competing set of products or brands and at what hierarchical level do various alternatives compete with one another. The tree approaches are particularly useful on this type of problem.

In addition to considering the specific product positioning questions the researcher wished to address and the relationship (s)he feels comfortable in assuming, the type of input data available or to be collected strongly influences the selection of the type of product questioning approach to be used.

Finally, the researcher must come to grips with the issue of heterogeneity in the perceptions of the population (sample). Each approach has methods for testing for heterogeneity and specific algorithms for properly dealing with various segmentation structures.

REFERENCES

1. Y. Wind and P. J. Robinson, Product positioning: an application of multidimensional scaling. In *Attitude Research in Transition* (Edited by R. I. Haley). American Marketing Association, Chicago, Ill. (1972).
2. J. Huber and M. B. Holbrook, Using attribute ratings for product positioning: some distinctions among compositional approaches. *J. Mktg Res.* **16,** 507–515 (1979).
3. J. R. Hauser and F. S. Koppleman, Alternative perceptual mapping techniques: relative accuracy and usefulness. *J. Mktg Res.* **16,** 495–506 (1979).
4. W. R. Dillon, D. G. Frederick and V. Tangparichdee, Decision issues in binding perceptual product spaces with multiattribute rating data. *J. Consum. Res.* **12,** 47–63 (1985).
5. Y. Wind, *Product Policy: Concepts, Methods, and Strategy.* Addison-Wesley, Reading, Mass. (1982).
6. K. Lancaster, *Consumer Demand: a New Approach.* Columbia Univ. Press, New York (1971).
7. P. E. Green and Y. Wind, *Multiattribute Decisions in Marketing*, Dryden Press, Hinsdale, Ill. (1973).
8. J. D. Carroll, Individual differences in multidimensional scaling. In *Multidimensional Scaling*; Vol. 1, *Theory* (Edited by R. N. Shephard, A. K. Romney and S. B. Nerlove). Seminar Press, New York (1972).
9. G. S. Day, A. D. Shoeker and R. Srivastava, Identifying competitive product markets: a review of customer oriented approaches. Working paper, Graduate School of Business, Univ. of Pittsburgh, Pa (1978).
10. J. A. Howard and J. N. Sheth, *The Theory of Buyer Behavior.* Wiley, New York (1969).
11. R. M. Johnson, *Multiple Discriminant Analysis Application to Marketing Research.* Market Facts Inc., Chicago, Ill. (1970).
12. P. E. Green and J. D. Caroll, *Mathematical Tools for Applied Multivariate Analysis.* Academic Press, New York (1976).
13. W. R. Dillon and M. Goldstein, *Multivarariate Analysis: Methods and Application.* Wiley, New York (1984).
14. J. Johnston, *Econometric Methods*, 2nd edn. McGraw-Hill, New York (1972).
15. R. A. Harshman and E. L. Margaret, Data processing and the extended PARFAG model. In *Research Methods Multi-mode Data Analysis* (Edited by H. G. Lawer *et al.*), pp. 216–284. Praeger, New York (1984).
16. P. E. Green and V. R. Rao, *Applied Multidimensional Scaling: a Comparison of Approaches and Algorithms*, pp. 65–68. Holt, Rinehart & Winston, New York (1972).
17. M. B. Holbrook and J. Huber, Separating perceptual dimensions from affective overtomes: an application to consumer aesthetics. *J. Consum. Res.* **5,** 272–283 (1979).

18. P. E. Green, Y. Wind and H. J. Claycamp, Brand features congruence mapping. *J. Mktg Res.* **12,** 306–313 (1975).
19. P. E. Green and F. J. Carmone, *Multidimensional Scaling and Related Techniques in Marketing Analysis*. Allyn & Bacon, Boston, Mass. (1970).
20. W. DeSarbo and V. R. Rao, A constrained methodology for product positioning. *Mtkg Sci.* 1–19 (1986).
21. P. E. Green and V. R. Rao, Configuration synthesis in multidimensional scaling. *J. Mtkg Res.* **9,** (1972).
22. S. Albers and K. Brockhoff, A procedure for new product positioning in an attribute space. *Eur. J. opl Res.* **1,** 230–238 (1977).
23. G. L. Urban, PERCEPTOR: a model for product positioning. *Mgmt Sci.* **21,** 858–871 (1975).
24. P. E. Green, J. D. Carroll and S. M. Goldberg, A general approach to product design optimization via conjoint analysis. *J. Mktg* **45,** 17–37 (1981).
25. J. J. Chang and J. D. Carroll. How to use MDPREF, a computer program for multidimensional analysis of preference data. Bell Telephone Labs, Murray Hill, N.J. (1969).
26. J. D. Carroll and J. J. Chang, A new method for dealing with individual differences in multidimensional scaling. *Mimeographed.* Bell Telephone Labs, Murray Hill, N.J. (1969).
27. J. D. Carroll and J. J. Chang, Relating preference data to multidimensional scaling solutions via a generalization of Coombs' unfolding model. Bell Telephone Labs, Murray Hill, N.J. (1967).
28. P. E. Green and J. McMennamin, Marketing position analysis. In *Marketing Managers' Handbook* (Edited by S. H. Britt). The Dartnell Corp., Chicago, Ill. (1973).
29. G. L. Urbana and J. R. Hauser, *Design and Marketing of New Products*. Prentice-Hall, Englewood Cliffs, N.J. (1980).
30. J. W. Keon, Product positioning: TRINODAL mapping of brand images, and images and consumer preference. *J. Mktg Res.* **20,** 380–392 (1983).
31. W. L. Wilkie and E. A. Pessemier, Issues in marketing's use of multiattribute models. *J. Mktg Res.* **10,** 428–441 (1973).
32. P. E. Green and V. Srinivasan, Conjoint analysis in consumer research: issues and outlook. *J. Consum. Res.* **5,** 103–123 (1978).
33. R. M. Johnson, Tradeoff analysis of consumer values. *J. Mktg Res.* **11,** 121–127 (1974).
34. J. B. Kruskal, Analysis of factorial experiments by estimating monotone transformations of the data. *J. R. statist. Soc. Ser. B.* **27,** 251–263 (1965).
35. V. Srinivasan and A. D. Shocker, Linear programming techniques for multidimensional analysis of preferences. *Psychometrika* **38,** 337–370 (1973).
36. R. J. Johnson, Varieties of conjoint measurement. Working paper, Market Facts Inc., Chicago, Ill. (1973).
37. D. McFadden, Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics* (Edited by P. Zerembka), pp. 105–142. Academic Press, New York (1973).
38. V. R. Rao and F. Winter, Application of the multivariate probit model for market segmentation and product design. Working paper 388, Graduate School of Business Administration, Univ. of Illinois at Urbana (1977).
39. V. R. Rao and D. J. Sabavala, Inference of hierarchical choice process from panel data. *J. Consum. Res.* **8,** 85–96 (1981).
40. R. K. Srivastva, R. P. Leone and A. D. Shocker, Market structure analysis: Hierarchical clustering of products based on substitution-in-use. *J. Mktg* **45,** 38–48 (1981).
41. M. U. Kalwani and D. G. Morrison, A parsimonious description of the Hendry System. *Mgmt Sci.* **23,** 467–477 (1977).
42. R. N. Shepard and P. Arabie, Additive clustering representation of similarities as combinations of discrete overlapping similarities. *Psychol. Rev.* **86,** 87–123 (1979).
43. S. Sattath and A. Tversky, Additive similarity trees. *Psychometrika* **42,** 319–345 (1977).
44. J. Corter and A. Tversky, Extended similarity trees. Unpublished paper, Dept of Psychology, Stanford Univ., Calif. (1983).
45. A. J. Phipps, J. Douglas Carroll, W. DeSarbo and J. Wind, Overlapping clustering: a new method for product positioning. *J. Mktg Res.* **18,** 310–317 (1981).
46. P. Arabie, J. D. Carroll, W. DeSarvo and J. Wind, Overlapping clustering—a new method for product positioning. *J. Mktg Res.* **17,** 310–317 (1981).
47. P. Arabie and J. D. Carroll, MAPCLUS: a mathematical programming approach to fitting the ADCLUS mode. *Psychometrika* **45,** 211–235 (1980).
48. A. D. Shocker and V. Srinivasan, Multiattribute approaches for product concept evaluation and generation: a critical review. *J. Mktg Res.* **16,** 159–180 (1979).
49. A. D. Shocker and V. Srinivasan, A consumer based methodology for the identification of new product ideas. *Mgmt Sci.* **21,** 858–871 (1974).
50. J. R. Hauser and G. L. Urban, A normative methodology for modeling consumer response to innovation. *Ops Res.* **25,** 579–619 (1977).
51. P. E. Green and Y. Wind, New way to measure consumer judgments. *Harvard Bus. Rev.* **53,** 107–117 (1975).
52. D. R. Wittink and D. Montgomery, Predictive validity of tradeoff analysis for alternative segmentation scheme. In *1979 AMA Educators' Conf. Proc.* (Edited by N. Beckwith *et al.*). American Marketing Association, Chicago, Ill. (1979).
53. W. L. Moore, Levels of aggregation in conjoint analysis: an empirical comparison. *J. Mktg Res.* **17,** 516–523 (1980).
54. A. A. Kuehn and R. L. Day, Strategy of product quality. *Harvard Bus. Rev.* **78,** 151–169 (1962).
55. J. C. Huber and W. L. Moore, A comparison of alternative ways to aggregate individual conjoint analyses. In *1979 AMA Educators' Conf. Proc.* (Edited by N. Beckwith *et al.*), pp. 64–68. American Marketing Association, Chicago, Ill. (1979).
56. R. J. Calatone and A. G. Sawyer, Stability of benefit segments. *J. Mktg Res.* **15,** 395–404 (1978).
57. E. A. Pessemier, Market structure analysis for new product and market opportunities. *J. contemp. Bus.* **4,** 35–67 (1975).
58. Y. Wind, Issues and advances in segmentation research. *J. Mktg Res.* **15,** 317–337 (1978).
59. R. G. Montanelli Jr and L. G. Humphreys, Latent roots of random data correlation matrices with squared multiple correlations on diagonal: a Monte Carlo study. *Psychometrika* **41,** 341–348 (1976).
60. J. D. Carroll and J. J. Chang, How to use INDSCAL. A computer program for canonical decomposition of *n*-way tables and individual differences in multidimensional scaling. Bell Telephone Labs, Murray Hill, N.J. (1969).

61. F. S. Zufryden, A zero–one integer programming model for market segmentation and product positioning. *J. Opl Res. Soc.* **30,** 63–70 (1979).
62. B. Gavish, D. Horsky and K. Srikant, Optimal positioning of a new product. *Mgmt Sci.* **29,** 1227–1297 (1983).
63. J. H. May, A. D. Shocker and D. Sudharshan, A simulation comparison of methods for new product location. Working paper, Graduate School of Business, Univ. of Pittsburgh, Pa (1982).
64. B. R. Parker and V. Srinivasan, A consumer preference approach to the planning of rural primary health care facilities. *Ops Res.* **24,** 991–1025 (1976).
65. J. R. Hauser and S. M. Shugan, Efficient measurement of consumer preference functions: a general theory for intensity of preference. Working paper 602-001, Graduate School of Management, Northwestern Univ., Evanston, Ill. (1977).
66. G. N. Punj and R. Staelin, The choice process for graduate business schools. *J. Mktg Res.* **15,** 588–598 (1978).
67. J. E. Scott and P. Wright, Modeling, an organizational buyer's product evaluation strategy: validity and procedural considerations. *J. Mktg Res.* **13,** 211–244 (1976).
68. C. Fraser and J. W. Bradford, Competitive market structure analysis: Principal partitioning of revealed substitutabilities. *J. Consum. Res.* **10,** 15–29 (1983).
69. G. S. Day, A. D. Shocker and R. K. Srivastava, Customer oriented approaches to identifying product-markets. *J. Mktg* **43,** 8–19 (1979).