Perceptual mapping has been used extensively in marketing. This powerful technique is used in new product design, advertising, retail location, and many other marketing applications where the manager wants to know (1) the basic cognitive dimensions consumers use to evaluate "products" in the category being investigated and (2) the relative "positions" of present and potential products with respect to those dimensions. For example, Green and Wind (1973) use similarity scaling to identify the basic dimensions used in conjoint analysis. Pessemier (1977) applies discriminant analysis to produce the joint-space maps that are used in his DESIGNR model for new product design. Hauser and Urban (1977) use factor analysis to identify consumer perceptions and innovation opportunities in their method for modeling consumer response to innovation. All of these researchers report empirical applications in a number of product and service categories. When used correctly perceptual mapping can identify opportunities, enhance creativity, and direct marketing strategy to the areas of investigation most likely to appeal to consumers.

Perceptual mapping has received much attention in the literature. Though varied in scope and application, this attention has been focused on refinements of the techniques, comparison of alternative ways to use the techniques, or application of the techniques to marketing problems. Few direct comparisons have been made of the three major techniques—similarity scaling, factor analysis, and discriminant analysis. In fact, most of the interest has been in similarity scaling because of the assumption that similarity measures are more accurate measures of perception than direct attribute ratings despite the fact that similarity techniques are more difficult and more expensive to use than factor or discriminant analyses.

In practice, a market researcher has neither the time nor the money to simultaneously apply all three techniques. He/she usually selects one method and uses it to address a particular marketing problem. The market researcher must decide whether the added insight from similarity scaling is worth the added expense in data collection and analysis. Furthermore, if the market researcher selects an attribute-based method such as factor analysis or discriminant analysis he/she wants to know which method is better for perceptual mapping and how such maps compare with...
those from similarity scaling. To answer these questions, one must compare the alternative mapping techniques.

One way to compare these techniques is theoretically. Each technique has theoretical strengths and weaknesses and the choice of technique depends on how consumers actually react to the alternative measurement tasks. Theoretical hypotheses are established in the next section.

Another comparison approach is Monté Carlo simulation. This useful investigative tool has been employed by researchers to explore variations in similarity scaling and other techniques (Carmone, Green and Jain 1978; Cattin and Wittink 1976; Pekelman and Sen 1977). In perceptual mapping Monté Carlo simulation can compare the ability of various techniques to reproduce a hypothesized perceptual map, but it requires that the researcher assume a basic cognitive structure of the individual. Monté Carlo simulation leaves unanswered the empirical question of whether the analytic technique can adequately describe and predict an actual consumer’s cognitive structure.

The comparison procedure we use is practical and is based on theoretical arguments and empirical analyses to identify which procedures yield results most useful for marketing research decisions. If the theoretical arguments are supported, researchers can continue to subject a technique to empirical tests in alternative product categories. In this way, one gains insight about the techniques by learning their strengths and weaknesses. If and when the hypotheses are falsified new theories will emerge.

We chose the following guidelines for the comparison.

1. The marketing research environment should be representative of the way the techniques are used empirically.
2. The sample size and data collection should be large enough to avoid exploiting random occurrences and should have no relative bias in favor of the techniques identified as superior.
3. The use of the techniques should parallel as closely as possible the recommended and common usage.
4. The criteria of evaluation should have managerial and research relevance.

To fulfill these criteria, we chose an estimation sample and a saved data sample of 500 consumers each, drawn from residents of Chicago’s northern suburbs. The application area is perceptions of the attractiveness of shopping areas in the northern suburbs and the criteria are the ability to predict consumer preference and choice, interpretability of the solutions, and ease of use.

**HYPOTHESES: THEORETICAL COMPARISON**

We begin with a brief synopsis of the techniques. Because this synopsis is not a complete mathematical description, we have indicated the appropriate references.

**Similarity scaling** develops measures of consumers’ perceptions from consumer judgments with respect to the relative similarity between pairs of products. Though consumers are asked to judge similarity between products, the definition of similarity usually is left unspecified. The statistical techniques select relative values for two, three, or four perceptual dimensions such that distance between products best corresponds to measured similarity. Green and Rao (1972) and Green and Wind (1973) provide mathematical details.

For empirical studies with large sample sizes, common space representations are developed such that each consumer’s i perception $\hat{x}_{ip}$ of product j along dimension d is a “stretching” of the common representation $\bar{x}_{jd}$. The technique, INDSCAL (Carroll and Chang 1970), simultaneously estimates the common space $\bar{x}_{jd}$ for all i and estimates individual weights $v_{id}$ to “stretch” these dimensions. Effectively, $\hat{x}_{ipd} = v_{ipd}^1/2 \bar{x}_{jd}$. For details see Carroll and Chang (1970).

The dimensions are named by judgment or by a regression-like procedure called PROFIT (Carroll and Chang 1964). In PROFIT, consumers are asked to rate each product on specific attributes, e.g., “atmosphere” for shopping centers. The ratings are dependent variables in a regression (possibly monotonic) on the perception measures, $\hat{x}_{ipd}$, which serve as explanatory variables. The regression weights, called directional cosines, indicate how strongly each perception measure relates to each attribute rating. For details see Carroll and Chang (1964).

**Factor analysis** begins with the attribute ratings. The assumption is that there are really a few basic perceptual dimensions, $\bar{x}_{jd}$. Many of the attribute ratings are related to each perceptual dimension. Factor analysis examines the correlations among the attributes to identify these basic dimensions. For statistical details see Harmon (1967) and Rummel (1970). Because concern is with the basic structure of perception, the correlations between attribute ratings are computed across products and consumers (sum over i and j). The perceptions of products are measured by “factor scores” which are based on the attribute ratings. The dimensions are named by examining “factor loadings” which are estimates of the correlations between attribute ratings and perception measures. In applications, attribute ratings are first standardized by individual to minimize scale bias.

**Discriminant analysis** also begins with the attribute ratings, but rather than examining the structure of attribute correlations, discriminant analysis selects the (linear) combinations of attributes that best discriminate between products. See Cooley and Lohnes (1971) and Johnson (1970) for mathematical details. Because concern is with the ability to differentiate products,
the dependent measure is “product rated” and the explanatory variables are the attribute ratings. The analysis is run across consumers to find a common structure. The perceptions are measured by “discriminant scores” which are estimates, based on the attribute ratings, of the perceptual dimensions, $x_{jd}$, that best distinguish products. The dimensions are named by examining “discriminant scores” which are the weightings of the attributes that make up a discriminant dimension or by computing correlations that are equivalent to factor loadings. In applications, the discriminant dimensions are constrained to be uncorrelated (also called orthogonal).

**Comparison of Similarity Scaling and Attribute-Based Techniques**

A major difference between similarity scaling and the attribute-based techniques is the consumer task from which the perceptual measures are derived. Attribute ratings are more direct measures of perceptions than similarity judgments, but may be incomplete if the set of ratings is not carefully developed. Similarity judgments introduce an intermediate construct (similarity) but the judgments are made with respect to the actual product rather than specific attribute scales. A priori, if the set of attributes is relatively complete there is no theoretical reason to favor one measure over the other.

Another difference is the treatment of variation among consumers. In the attribute-based techniques a common structure is assumed, but the values of individual measures ($x_{jd}'$ or $x_{jd}$) are not restricted. In similarity scaling (INDSCAL) $x_{jd}$ is restricted to be at most a stretching of the common measure, $x_{jd}'$. This means that although complete reversals are allowed, no other change in rank order is allowed. For example, INDSCAL will not allow one consumer to evaluate relative sweetness in the order Pepsi, Coke, Royal Crown while another evaluates sweetness as Coke, Royal Crown, Pepsi. This restriction limits the applicability of the similarity scaling.

Finally, similarity scaling is limited by the number of products. At least seven or eight are needed for maps in two or three dimensions (Klahr 1969). There are no such restrictions for factor analysis. The restriction for discriminant analysis is the number of products minus one. This argument favors attribute-based techniques if the number of products in a consumer’s evoked set is small; it favors neither technique if the number of products is large. In practice, the evoked set averages about three products (Silk and Urban 1978).

On the basis of these arguments, if the attribute set is reasonably complete, attribute-based techniques should provide better measures of consumer perception than similarity scaling (as implemented by INDSCAL).

**Comparison of Factor Analysis and Discriminant Analysis**

Factor analysis is based on the correlations across consumers and products. Discriminant analysis is limited to dimensions that, on average, distinguish among products. Thus factor analysis should use more attributes than discriminant analysis in the dimensions and therefore produce richer solutions. For example, consider Mercedes Benz and Rolls Royce. Suppose that the true perceptual dimensions are country of origin and reliability and that only reliability affects preference and choice. Suppose that perceptions of country of origin differ among products. Discriminant analysis will identify only country of origin. Factor analysis will identify both dimensions.

On the basis of this type of argument, one expects factor analysis to provide a richer perceptual structure than discriminant analysis. It should be able to use more of the attribute ratings and should identify perceptual dimensions that predict preference and choice better than discriminant analysis dimensions.

**EMPIRICAL SETTING: SHOPPING LOCATIONS**

The empirical setting is north suburban Chicago and the “product” category is overall image of shopping locations, an increasingly important area of investigation in marketing. From the perspective of retailers, shopping center managers, and community planners, the sensitivity of destination choice behavior to the image or attractiveness of the shopping location provides an important opportunity to develop strategies to attract shoppers.

In terms of an empirical comparison of perceptual techniques, shopping location choice provides a strong test of the model’s explanatory and predictive capabilities. Shopping location choice is a complex phenomenon, difficult to model and difficult to understand. If a perceptual technique does well in this category, it is likely to do well in a less complex category. (The complexity of the category should introduce no relative bias in the model comparisons.)

We develop models based on seven shopping areas including downtown Chicago and six suburban shopping centers of very different characteristics. The locations represent the types of shopping locations available to the residents of the suburbs north of Chicago including large, medium, and small shopping areas with exclusive, general merchandise, or discount orientations.

The data were collected by a self-administered questionnaire. (For details see Stopher, Watson, and Blinn 1974.) The attributes were measured by 16 five-point rating scales chosen in an attempt to get as
complete a list as possible without causing consumer wearout. (They were selected and refined on the basis of literature reviews, preliminary surveys, and qualitative research.) The similarity judgments were measured by pairwise comparisons on a seven-point scale. The respondent judged all pairs of the products in his/her evoked set (measured by a knowledge question). Table 1 contains example instructions for the attribute ratings and similarity judgments.

The dependent measures for predictive testing were obtained in the same survey. Choice was self-reported frequency of visits. Preference, also called "attractiveness," was rank-order preference in which availability/accessibility was held constant. (Pretests indicated consumers were comfortable with this task.) Availability/accessibility was measured as map distance from the consumer's residence to the shopping location. The model (explained in the next section) was drawn from the retailing and transportation literature and is given as follows.

This two-stage model allows managers and researchers to measure the "attractiveness" of a shopping center independently of its location. Thus one can readily investigate strategies, such as improving the atmosphere of a shopping center, that may be more cost effective than relocation strategies. This ability is especially useful when relocation is not an option as is the case in many downtown shopping areas.

In the original data set, 37,500 mailback questionnaires were distributed at four of the shopping locations. Of these, 6,000 were returned as complete and usable questionnaires. Although this low return rate may cause some nonresponse bias, the bias should not affect relative comparisons. Because similarity scaling requires at least seven stimuli, the data were screened to select those consumers who indicated

---

**Table 1**

**EXAMPLE QUESTIONS FOR ATTRIBUTE SCALES AND SIMILARITY JUDGMENTS**

In this question, we would like you to rate each of the shopping centers on these characteristics. We have provided a range from good to poor for each characteristic. We would like you to tell us where you feel each shopping center fits on this range.

For example:

- **Eating Facilities**
  - Good
  - Poor

---

**RATINGS OF THE ATTRIBUTES OF SHOPPING LOCATIONS**

Again, if all the shopping centers were equally easy to get to, how similar do you think they are to each other? In answering this question, please think about your preference to shop at them for the goods you came to buy. Check the box which best describes how similar they are. Please be sure to do this for all pairs of shopping centers.

<table>
<thead>
<tr>
<th>Woodfield and Chicago Loop</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edens Plaza and Golf Mill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Woodfield and Plaza del Lago</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SIMILARITY JUDGMENTS**
knowledge of and who rated all seven stimuli (1,600 consumers). Any bias introduced by this screening should favor similarity scaling and thus favor falsifying our hypotheses. Finally, 500 of these consumers were randomly selected for an estimation sample and 500 others for saved data testing.

One source of bias is that samples were taken at four locations but judgments were made for seven locations. This sampling technique should produce no bias in the relative predictive ability of the perceptual techniques but could bias the coefficients in the preference and choice models. Fortunately, this bias can be corrected with choice-based sampling (CBS) estimation procedures. Manski and Lerman (1977) give theoretical proofs and Lerman and Manski (1976) provide intuitive arguments. To test CBS empirically for the data, parallel analyses were performed using only the four sampled shopping locations. Consistent with the theory, all models were statistically equivalent. Koppelman and Hauser (1977) give details of these comparisons.

EMPIRICAL PROCEDURE

So that other researchers can replicate the tests, we describe in this section how the measures were developed. In selecting procedures an attempt was made to follow common and recommended usage as closely as possible. When potential variations occurred (rotation vs. nonrotation, direct vs. indirect similarities, three vs. four dimensions) sensitivity tests were performed.

Similarity Scaling

Pairwise similarities were changed to rank order similarities and input to INDSCAL. Three dimensions were selected on the basis of an elbow in stress values. Four dimensions were not possible with the limited number of products.

Factor Analysis

The attributes were standardized by individual. Correlations were computed across consumers and products. Four dimensions were selected by using elbow and interpretability criteria on the common factor solution with varimax rotation. Limited testing was done with three dimensions for comparability to similarity scaling.

Discriminant Analysis

The attributes were standardized by individual. A discriminant analysis was run with the product rated as the dependent variable and the 16 attributes as explanatory variables. The elbow rule and significance statistics suggested at least three dimensions (97.5% of trace), possibly four because the chi square test was still significant. Most analyses were done with four dimensions although limited analyses were done with three dimensions. Varimax rotation improved the

interprecations slightly. Dimensions are constrained to be orthogonal.

Preference

The linear compensatory form was chosen because of its widespread use in marketing (Wilkie and Pessier 1973), because Monté Carlo simulation has shown it reasonably representative of more complex forms such as disjunctive, conjunctive, and additive (Carmon, Green, and Jain 1978), and because many of the nonlinear models such as conjoint analysis (Green and Srinivasan 1978, Green and Wind 1975) require (more extensive) personal interviews rather than the mailback format used by Stopher, Watson, and Blin (1974).

Mathematically, this model is given by:

\[ p_{ij} \sim \sum_{i=1}^{m} w_i d_{ij}, \]

where \( p_{ij} \) is individual \( i \)'s rank for product \( j \), \( \sim \) indicates monotonic, \( d_{ij} \) is \( i \)'s perception of \( j \) along the \( l \)th dimension, and \( w_i \) is the importance weight of the \( l \)th dimension. (For similarity scaling \( d_{ij} = \bar{x}_{i\delta} \), for factor analysis \( d_{ij} = x_{i\delta} \), for discriminant analysis \( d_{ij} = x_{i\delta} \).

Because there is a possibility that any single preference estimation procedure for the \( w_i \)'s will favor one or another perceptual model, two estimation procedures were simultaneously tested: preference regression and first preference logit. Preference regression is a metric technique which replaces monotonicity (\( \sim \)) by equality (=) and uses ordinary least squares with equation 1. First preference logit is a monotonic technique based on estimating the probability that a consumer will rank a product as first preference. In logit, maximum likelihood techniques are used to determine \( w_i \). For a more complete description, see McFadden (1970). In both models, choice-based sampling variables are used to provide consistent estimates of \( w_i \) (Manski and Lerman 1977).

Choice

The multinomial logit model is used to predict choice. (The logit model is based on a probabilistic interpretation of choice. For estimation and prediction equations, see McFadden 1970.) The dependent variable is frequency of visits and the explanatory variables are distance and estimated preference, \( \bar{p}_{ij} \), where \( \bar{p}_{ij} \) is generated by equation 1 and the estimates, \( \hat{w}_i \), of the importance weights. That is, \( \bar{p}_{ij} = \hat{w}_i d_{ij} \).

Finally, we chose to test the assumption of a two-stage model by using a revealed preference logit model. The dependent variable was frequency and the explanatory variables were distance and the perceptual measures, \( d_{ij} \). Choice-based sampling variables were included for consistent estimates.
Saved Data Predictions

One begins with the standardized attribute ratings and similarity judgments from the saved data sample. Factor score coefficients from the estimation sample and the attribute ratings from the saved data are used to create factor scores, \( x_{\mu\delta} \) (see Rummel 1970). Discriminant score coefficients from the estimation sample and the attribute ratings from the saved data are used to create discriminant scores, \( x_{\mu\delta}' \) (see Cooley and Lohnes 1971). INDSCAL is run on the saved data similarity judgments to create similarity scores, \( x_{\mu\delta}'' \). If anything, this procedure should favor similarity scaling. Indeed, an alternative procedure of generating the similarity scores from the attribute ratings provided poorer predictive results for similarity scaling and thus is even stronger support for the hypotheses.

For each combination of models (perception, preference, choice) the estimated importance weights, \( w_{ij} \), from the estimation sample are applied to the \( d_{ij} \)’s to create \( \hat{p}_{ij} \). These are rank ordered and compared with the actual \( p_{ij} \). The estimated relative weights of \( \hat{p}_{ij} \) and distance then are used (with the relevant logit equation) to estimate the relative frequency of visits. These predictions are compared with reported frequency of visits.

Predictive Tests

Preference prediction was measured by the percentage of consumers who ranked first the shopping center that the model predicted as first. This measure is straightforward and is commonly reported in the literature.

Choice prediction is more difficult to measure. First, the models predict the probability of choice which must be compared with frequency of choice. This problem is resolved by using percentage uncertainty which is an information theoretic statistic that measures the percentage of uncertainty (entropy) explained by the probabilistic model (see Hauser 1978; McFadden 1970; Silk and Urban 1978). The second problem arises because both accessibility and the perceptual dimensions (through preference) are used to predict choice. Only the incremental gains in uncertainty due to the perceptual measures are of interest. Because the information measure is additive, we use the gain in percentage uncertainty achieved by adding the perceptual dimensions to a model based on distance alone.

Other measures including rank order preference recovery, percentage of consumers choosing the maximum probability shopping location, and mean absolute error in market share prediction are not reported here because each of these measures ranked the perceptual model in the same relative order as do the reported measures. For these statistics see Koppelman and Hauser (1977).

Predictive Tests

The results of the predictive tests are reported in Table 2. Preference recovery was not computed for revealed preference logit which is estimated directly on choice. Table 2 also reports preference and choice prediction for models using distance and the 16 standardized attribute scales as explanatory variables. These statistics are computed to examine whether the perceptual dimensions provide more or less information than a full set of attribute ratings.

To better understand Table 2, it is useful to consider some base level values. The maximum value for preference recovery is 100%, but most empirical models do not obtain values even near 100%. Rather these measures should be compared with that attainable by purely random assignment, i.e., all locations equally likely to be chosen, and that attainable by assigning consumers to shopping centers in proportion to market share. These values are 14.3% preference recovery for the equally likely model and 26.7% preference recovery for the market share model. Rigorous statistical tests are not applicable in comparing one perceptual technique with another because the preference models are based on different explanatory variables. But, intuitively, differences in preference recovery can be compared via the maximum standard deviation, 2.2, which would result under the assumption that for a given model each observation (consumer) has an equal probability of being correctly classified.

The significance of the choice models can be tested because the uncertainty explained is proportional to a chi square statistic (Hauser 1978). All models are significant at the .01 level. The chi square’s cutoff, which is less than 0.4, can also be taken as an intuitive measure in comparing the alternative perceptual techniques.

First examine the predictive tests. Factor analysis is superior to similarity scaling and discriminant analysis for all preference models and both prediction measures. This evidence supports the hypotheses. These results are particularly significant because any incompleteness in the attribute scales would favor similarity scaling over discriminant analysis and factor analysis. Furthermore, the sample was screened to favor similarity scaling.

The fact that factor analysis does well in comparison with models based on attribute scales indicates that...
very little information is lost by using the reduced factors rather than the full set of attributes. The poor showing of preference regression on the attribute scales is the result of multicollinearity. The poor showing of discriminant analysis in relation to the attribute scales supports the hypothesis that concentrating on variations between products neglects dimensions that are important in preference and choice.

Next, examine the saved data test. Factor analysis is still superior to both discriminant analysis and similarity scaling. Both of the attribute-based measures hold up well. The small improvement in the statistics may be attributed to random variation between data sets. Similarity scaling predicts poorly on saved data, doing worse than the naïve model which assigns consumers proportional to market share (26.7% preference recovery). The model does so poorly in relation to distance alone that it adds large amounts of uncertainty (negative uncertainty explained) in choice prediction.

Thus the saved data tests support the hypotheses. Table 2 is one empirical comparison. Replications, sensitivity, and threats to validity are examined in the next section, but first a more qualitative analysis is used to determine whether these comparisons are consistent with managerial interpretations of the techniques.

Interpretability

First examine how the perceptual dimensions from each technique relate to the attribute scales. These relationships are reported in Table 3. The underlined numbers (directional cosines, factor loadings, or discriminant coefficients) indicate a strong relationship between the underlined attribute (row) and the dimension (column). The dimension names are composites of the underlined attributes.

Despite marked superficial similarities, the different models demonstrate striking differences in interpretation. Factor analysis uses all the attribute scales whereas discriminant analysis uses only 10 of the 16 scales, and only five of those 10 have discriminant scores greater than 0.5. This outcome is consistent with the theoretical hypothesis that concentrating on variance between products neglects information. Note also that the third discriminant dimension, “value vs. satisfaction,” is a mixed dimension positively related to value but negatively related to satisfaction. Compare this to the factor analysis solution which has no positive/negative mixed dimension. This difference probably arises because value and satisfaction may be negatively correlated among existing shopping locations (discriminant analysis) but not in the way consumers rate these locations (factor analysis).

These differences are important to the manager who wants as much control over the market as possible. For example, many of the variables left out of the discriminant solution (e.g., “layout of the store,” “return and service,” “sales assistants”) could be affected by relatively low-cost changes in shopping center operation. Because discriminant analysis uses fewer variables the manager can analyze fewer strate-
Table 3  
STRUCTURAL COMPARISON OF PERCEPTUAL MODELS

<table>
<thead>
<tr>
<th>Fundamental attributes</th>
<th>Variety</th>
<th>Quality and satisfaction</th>
<th>Value</th>
<th>Parking</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Layout of store</td>
<td>.267</td>
<td>.538</td>
<td>.156</td>
<td>.200</td>
</tr>
<tr>
<td>2. Return and service</td>
<td>.095</td>
<td>.528</td>
<td>.343</td>
<td>.255</td>
</tr>
<tr>
<td>3. Prestige of store</td>
<td>.338</td>
<td>.822</td>
<td>-.001</td>
<td>-.058</td>
</tr>
<tr>
<td>4. Variety of</td>
<td>.665</td>
<td>.327</td>
<td>.309</td>
<td>-.185</td>
</tr>
<tr>
<td>merchandise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Quality of</td>
<td>.327</td>
<td>.817</td>
<td>.037</td>
<td>-.074</td>
</tr>
<tr>
<td>merchandise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Availability of</td>
<td>.159</td>
<td>.337</td>
<td>.487</td>
<td>.049</td>
</tr>
<tr>
<td>credit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Reasonable price</td>
<td>.067</td>
<td>-.063</td>
<td>.599</td>
<td>.113</td>
</tr>
<tr>
<td>8. “Specials”</td>
<td>.223</td>
<td>.074</td>
<td>.739</td>
<td>.008</td>
</tr>
<tr>
<td>9. Free parking</td>
<td>-.15</td>
<td>.068</td>
<td>.043</td>
<td>.811</td>
</tr>
<tr>
<td>10. Center layout</td>
<td>.03</td>
<td>.308</td>
<td>.074</td>
<td>-.560</td>
</tr>
<tr>
<td>11. Store atmosphere</td>
<td>.080</td>
<td>.658</td>
<td>.034</td>
<td>.400</td>
</tr>
<tr>
<td>12. Parking available</td>
<td>.145</td>
<td>.105</td>
<td>.108</td>
<td>.841</td>
</tr>
<tr>
<td>13. Center atmosphere</td>
<td>.244</td>
<td>.694</td>
<td>.040</td>
<td>.404</td>
</tr>
<tr>
<td>14. Sales assistants</td>
<td>.173</td>
<td>.561</td>
<td>.147</td>
<td>.319</td>
</tr>
<tr>
<td>15. Store availability</td>
<td>.619</td>
<td>.320</td>
<td>.204</td>
<td>.034</td>
</tr>
<tr>
<td>16. Variety of stores</td>
<td>.829</td>
<td>.288</td>
<td>.160</td>
<td>-.173</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fundamental attributes</th>
<th>Variety</th>
<th>Quality and satisfaction</th>
<th>Value vs. satisfaction</th>
<th>Parking and satisfaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Layout of store</td>
<td>.217</td>
<td>.497</td>
<td>.640</td>
<td>502</td>
</tr>
<tr>
<td>2. Return and service</td>
<td>.318</td>
<td>.122</td>
<td>.940</td>
<td>502</td>
</tr>
<tr>
<td>3. Prestige of store</td>
<td>.297</td>
<td>.804</td>
<td>.515</td>
<td>502</td>
</tr>
<tr>
<td>4. Variety of</td>
<td>.929</td>
<td>.360</td>
<td>.084</td>
<td>502</td>
</tr>
<tr>
<td>merchandise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Quality of</td>
<td>.295</td>
<td>.811</td>
<td>-.505</td>
<td>502</td>
</tr>
<tr>
<td>merchandise</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Availability of</td>
<td>.880</td>
<td>-.085</td>
<td>-.468</td>
<td>502</td>
</tr>
<tr>
<td>credit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Reasonable price</td>
<td>.485</td>
<td>-.853</td>
<td>.192</td>
<td>502</td>
</tr>
<tr>
<td>8. “Specials”</td>
<td>.786</td>
<td>-.594</td>
<td>.173</td>
<td>502</td>
</tr>
<tr>
<td>9. Free parking</td>
<td>-.294</td>
<td>-.550</td>
<td>.782</td>
<td>502</td>
</tr>
<tr>
<td>10. Center layout</td>
<td>-.447</td>
<td>.036</td>
<td>.894</td>
<td>502</td>
</tr>
<tr>
<td>11. Store atmosphere</td>
<td>-.199</td>
<td>.452</td>
<td>.869</td>
<td>502</td>
</tr>
<tr>
<td>12. Parking available</td>
<td>-.463</td>
<td>-.478</td>
<td>.747</td>
<td>502</td>
</tr>
<tr>
<td>13. Center atmosphere</td>
<td>-.099</td>
<td>.480</td>
<td>.872</td>
<td>502</td>
</tr>
<tr>
<td>14. Sales assistants</td>
<td>-.052</td>
<td>.411</td>
<td>.910</td>
<td>502</td>
</tr>
<tr>
<td>15. Store availability</td>
<td>.872</td>
<td>.429</td>
<td>-.236</td>
<td>502</td>
</tr>
<tr>
<td>16. Variety of stores</td>
<td>.921</td>
<td>.385</td>
<td>-.054</td>
<td>502</td>
</tr>
</tbody>
</table>

Comparing factor analysis with similarity scaling, one again sees a mixed dimension, "quality vs. value," but the similarity scaling solution uses all 16 attributes. This mixed dimension probably results from the INDSCAL assumption which concentrates on differences between products at the expense of variations in consumer perceptions. But in this case the "quality vs. value" is easier to understand and operate on than "satisfaction vs. value."

Thus, of the three models, factor analysis has the most managerially useful structure. Because all models have strong face validity and all use the basic five constructs—variety, quality, satisfaction, value, and
parking—we believe that the predictive and saved data superiority of factor analysis is due to better structure in identifying dimensions.

The second test of interpretability is the visual maps produced by each method (see Figure 1). These maps help managers identify how each shopping location is “positioned” in the marketplace. Thus a manager can know the relative strengths and weaknesses of each shopping location and can identify opportunities in the market.

Careful inspection of Figure 1 shows consistency among the models when the measured constructs are the same. For example, note the low scores for Korvette City on quality (satisfaction) and for Chicago Loop on parking or the high scores for Woodfield on variety and for Plaza del Lago on quality (satisfaction). These and many other “positionings” have strong face validity and are consistent with prior beliefs about the images of the seven shopping locations. But there is an important difference. Factor analysis is better for strategy development because it separates the dimensions in such a way that ambiguous interpretations (“quality vs. value” and “satisfaction vs. value”) are avoided.

**Ease of Use**

A final consideration is the investment in analysis time required by the techniques. Managers and researchers may be willing to accept more approximate models if the approximations allow great cost savings. Our experience indicates that the attribute-based techniques are relatively easy to use whereas similarity scaling is somewhat more difficult with respect to both data collection and data analysis. Both factor and discriminant analyses use only the attribute ratings whereas similarity scaling requires, in addition, judged similarities which can be a difficult consumer task. Once the data have been collected, factor and discriminant analyses are readily available on standard statistical packages (e.g., BMDP, SPSS), cost about $10-20 to run, and require little professional time. In sharp contrast, the special programs for similarity scaling require many exploratory runs, special FORTRAN programs for data transfer, and a series of statistical manipulations and data handling to develop a common space, estimate individual weights, and compute directional cosines. A single set of runs costs about $40 in computer time, but because various starting configurations and dimensions must be checked, the effective cost is about $150. Though the added direct computer costs are acceptable, the programming and analysis require significant professional time and could be very costly to organizations not familiar with the programs.

**Costs vary by computer. These costs represent $510 per cpu hour on a CDC 6400. See Dixon (1975) for BMDP and Nie et al. (1975) for SPSS.**
Table 4
SUMMARY OF THE RESULTS OF EMPIRICAL COMPARISON*

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theory</td>
<td>Factor analysis &gt; Discriminant analysis ~ Similarity scaling</td>
</tr>
<tr>
<td>Goodness of fit</td>
<td>Factor analysis &gt; Discriminant analysis ~ Similarity scaling</td>
</tr>
<tr>
<td>Prediction (saved data)</td>
<td>Factor analysis &gt; Discriminant analysis ~ Similarity scaling</td>
</tr>
<tr>
<td>Interpretation</td>
<td>Factor analysis ≥ Discriminant analysis ~ Similarity scaling</td>
</tr>
<tr>
<td>Ease of use</td>
<td>Factor analysis ~ Discriminant analysis ~ Similarity scaling</td>
</tr>
</tbody>
</table>

* > indicates superior, ≥ indicates probably superior, ~ indicates no major difference.

Summary

In this section representative attribute-based and similarity scaling techniques are compared to determine which perceptual mapping technique is most appropriate for marketing research applications. The empirical results, which support the hypotheses, suggest that the most popular technique, similarity scaling, may not be the best; rather, factor analysis may be better on predictability, interpretability, and ease of use (see Table 4). These results are subject to confirmation or qualification in other empirical applications but, at the very least, this empirical test raised issues worth further investigation by other marketing researchers.

THREATS TO VALIDITY, SENSITIVITY, AND REPLICATIONS

Empirical comparisons are difficult. Only by using the state of the art in each technique can one be fair to that technique. To attempt to extrapolate these comparisons, one must examine causes that threaten to make the results specific to the empirical sample.

Threats to Validity

Nonresponse bias, choice-based sampling, representativeness of the set of attributes, and screening on the seven stimuli in the evoked set have been discussed. They should introduce no relative bias favoring the hypotheses.

One threat to the hypotheses is potential halo effects in the attribute scales, i.e., the attribute scales may contain an affective as well as a cognitive component (Beckwith and Lehmann 1975). This threat is minimized by standardization of the attribute scales prior to analysis and by using attribute scales designed to measure only the cognitive component. Another threat is that the stimuli set contained at most seven products. Although this constraint is typical of real-world applications, it does suggest future comparisons with larger stimuli sets.

Sensitivity

The comparisons among techniques are found to remain consistent when limited changes are made in each technique. Rotation of discriminant solutions improved interpretability but not predictive ability (the rotated solution is reported). Use of indirect similarity measures did not improve either interpretability or predictive ability. The factor and discriminant analyses used four dimensions whereas similarity scaling was limited to three dimensions. Though this is a minor change in the degrees of freedom in the preference and choice models (490 vs. 489), it could have an effect. Limited testing showed that dropping the least significant dimension in the four-dimensional solutions caused very little shrinkage in prediction. (For example, with preference regression and discriminant analysis the shrinkage was about one-tenth of one percent in preference recovery.) Finally, use of alternative factor analysis solutions such as principal components or three-way analyses might improve prediction, but such improvement would only add support to the hypotheses.

Replications

Since the original study, the hypotheses have been tested on two other data sets. Using a sample of 120 graduate students, Simmie (1978) found preference recoveries of 67.4% for factor analysis, 51.9% for discriminant analysis, and 14.0% for similarity scaling. Simmie also found greater consistency across groups of students with factor analysis than with similarity scaling. Her product category was management schools. Using a random sample of Evanston residents, Englund, Hundt, and Lee (1978) found preference recoveries of 67.1% with factor analysis and 48.6% with discriminant analysis. Their product category was transportation mode choice (bus, walk, or car). The implications of both studies are consistent with our results.

IMPLICATIONS AND FUTURE RESEARCH

Perceptual mapping is an important marketing research tool used in new product planning, advertising development, product positioning, and many other areas of marketing. Strategies based on perceptual maps have led to increased profits, better market control, and more stable growth. Furthermore, much research is based on implications of market structure as identified by perceptual maps. Because of this
interest and use, it is crucial that the best mapping technique available be employed in these applications.

We provide and support hypotheses that factor analysis is superior to both similarity scaling and discriminant analysis for developing measures of consumer perceptions. In particular, factor analysis is likely to be superior in categories where:

1. The number of products in the average consumer's evoked set is relatively small (seven stimuli or less).
2. There is variation in the way consumers perceive products in the category.
3. Qualitative research has identified a set of attributes likely to represent the product category.

The presence or absence of these characteristics does not ensure the superiority of one technique, but without evidence to the contrary they can serve as guidelines.

The results of this one theoretical and empirical comparison are hoped to raise the issue of and the need for continued research to identify whether factor analysis is always superior or, if not, under what conditions the alternative mapping techniques should be used.

Confirmation of these comparisons awaits replication in other categories. In addition, further exploration may be appropriate for other preference models. For example, nonlinear models such as conjoint analysis should theoretically order the perceptual models in the same way, but empirical tests are warranted. We have used disaggregate preference models, i.e., models based on individual consumers ($p_i$, $d_{i,k}$ rather than $p_i$, $d_{i,k}$ where $i$ indexes consumers). In most applications these disaggregate models have proven superior to aggregate models which blur individual differences. Because aggregate models still are used in some marketing applications, these might be tested. For example, PREFMAP, which is a form of preference regression, can be used at both the aggregate (Pessemer 1977) and disaggregate level (Beckwith and Lehmann 1975). We hypothesize that because the superiority of factor analysis over discriminant analysis is based on individual differences, predictive comparisons might shift and discriminant analysis might do as well as factor analysis when analyses are limited to the aggregate level.

Other research might be of a more proactive nature, searching for improvements for the weaknesses of each technique. Theoretical developments might make it practical to relax the INDSCAL assumption. Further research could expand the similarity scaling solutions to more stimuli via use of concept statements. Segmentation on perceptions could be used prior to similarity scaling or discriminant analysis to ensure that there is little variation among consumers. These and other methodological developments are suggested by the examination and comparison of the alternative perceptual mapping techniques.

This area of comparative model development is important to marketing and deserves attention from marketing researchers.

REFERENCES


